EL DESPERTAR DE LAS MÁQUINAS PENSANTES

Una travesía por los entramados
de la Inteligencia Artificial
y sus infinitos reflejos

Gustavo de Elorza Feldborg

Elorza Feldborg, Gustavo de
El despertar de las máquinas pensantes : una travesía por los entramados de la Inteligencia Artificial y sus infinitos reflejos / Gustavo de Elorza Feldborg. - 1a ed - Mar del Plata : Eliana Inés Valzura, 2024.
Libro digital, EPUB

Archivo Digital: descarga y online
ISBN 978-631-00-4301-2

1. Educación. 2. Nuevas Tecnologías. 3. Inteligencia Artificial. I. Título.
CDD 371.3028

UNIVERSIDAD FASTA
FACULTAD DE CIENCIAS DE LA EDUCACIÓN

El despertar de las máquinas pensantes

Una travesía por los entramados de la inteligencia artificial y sus infinitos reflejos

UNIVERSIDAD FASTA
FACULTAD DE CIENCIAS DE LA EDUCACIÓN

El despertar de las máquinas pensantes

Una travesía por los entramados de la inteligencia artificial y sus infinitos reflejos

Gustavo de Elorza Feldborg

Edición general: Lic. Eliana Valzura. Ediciones Diapasón

Producción gráfica: Mg. María Eugenia Huinchulef
Secretaría de Investigación
Facultad de Ciencias de la Educación
Universidad FASTA

Diseño gráfico de cubierta: DG. Camila de Elorza

© 2024, Gustavo de Elorza Feldborg

Prohibida su reproducción total o parcial sin autorización del autor.

Elorza Feldborg, Gustavo de

El despertar de las máquinas pensantes : una travesía por los entramados de la Inteligencia Artificial y sus infinitos reflejos / Gustavo de Elorza Feldborg. - 1a ed - Mar del Plata : Eliana Inés Valzura, 2024.

ISBN 978-631-00-4301-2

A Dios, cuya guía y fortaleza han sido faro en mi travesía.
A mis hijos, espejos de mi ser,
en quienes hallo la eterna inspiración.
A aquellos invisibles arquitectos de este libro,
cuyas contribuciones son las verdaderas
piedras angulares de este libro.
Y a Caro, que con su genialidad me ayudó a
crear algo que resulta una oportunidad de aprender.

Índice

Capítulo I.
Cerebros Artificiales.
Ficciones y perspectivas de la Inteligencia Artificial en la vida de los humanos..... 3

Capítulo 2.
El destino algorítmico.
Trascendiendo los límites de la cuarta dimensión..... 13

Capítulo 3
Los Golem Digitales.
Prospectiva y perspectivas de la IA..... 32

Capítulo 4
Horizontes de la Inteligencia.
Desentrañando el significado detrás de la Inteligencia artificial y humana..... 51

Capítulo 5
Prometeo Digital. El nuevo fuego de los humanos..... 89

Capítulo 6
Anatomía de un *prompt.*
Elementos esenciales para una comunicación eficaz con la Inteligencia Artificial..... 102

Capítulo 7
¿Creando Monstruos?
La paradoja de la Inteligencia Artificial como el nuevo Frankenstein digital..... 163

Sobre el autor..... 200

Capítulo I

Cerebros Artificiales
Ficciones y perspectivas de la Inteligencia Artificial en la vida de los humanos

"Tlön, Uqbar, Orbis Tertius"[1], es un cuento de Jorge Luis Borges que explora sobre un mundo ficticio llamado Tlön y cómo este mundo se va infiltrando en la realidad. A partir de un encuentro con su amigo Bioy Casares, en el que este pronuncia una frase supuestamente leída en una enciclopedia, el narrador y su amigo comienzan una búsqueda de alguna referencia de ese lugar llamado Tlön que aborrece los espejos porque multiplican —como la cópula, dicen— las realidades. A medida que pasan los días se van enterando algo más sobre Tlön: que es una región de Uqbar —aunque no logran dar con ese sitio en ninguna enciclopedia—. Ese mundo, con su filosofía, su literatura, su cultura, su historia, es una invención de un

[1] Borges, J. L. (1997). *Ficciones*. Alianza editorial.

grupo de personas que está preparando una nueva enciclopedia, llamada *Orbis Tertius* ("tercer mundo"). Tal como una concatenación de espejos —imagen muy cara al imaginario borgeano— ese mundo tercero que ni siquiera es la primera imagen del real sino la narración de esa imagen, su descripción en palabras, poco a poco se va introduciendo en el mundo real subrepticiamente, borrando las fronteras entre uno y otro, hasta que el relato concluye con una conjetura inquietante: tal vez en cien años, aventura el narrador, habrá de cambiarlo por completo.

La inteligencia artificial, quizás un nuevo Tlön, un planeta digital construido con lenguajes de programación, ya opera hoy en nuestras vidas —aunque no lo sepamos o no advirtamos su alcance— y tiene un crecimiento exponencial en cuanto a sus capacidades.

La inteligencia artificial, quizás un nuevo Tlön, un planeta digital construido con lenguajes de programación, ya opera hoy en nuestras vidas —aunque no lo sepamos o no advirtamos su alcance— y tiene un crecimiento exponencial en cuanto a sus capacidades. La IA ha empezado un proceso de

"infiltración" en todas las áreas de nuestra vida cotidiana: los artefactos que hacen más sencilla la vida doméstica y la laboral, la agricultura, los servicios financieros, la automatización industrial, las comunicaciones y redes en las que interactuamos, las compras que realizamos, las operatorias bancarias, la medicina con la que se curan o previenen enfermedades, los juegos con los que nos entretenemos, los libros que leemos en línea, las consultas que hacemos en diversos organismos públicos o privados, los turnos que solicitamos para trámites y, por supuesto, la posibilidad de llevar con nosotros una IA interactiva en la palma de nuestra mano, complementando de forma exocerebral[2] nuestras capacidades cognitivas…

En fin: todo está permeado por la IA. No hace falta encerrarse en un laboratorio de robótica ni ser un experto en nuevas tecnologías. Esta situación, casi novedosa e impensada hace apenas unos años, nos

[2] El prefijo "exo" significa "fuera" o "exterior", adicionado a "cerebral" nos referimos, entonces, a algo que está más allá o fuera del cerebro.

plantea necesariamente nuevos desafíos y hasta hace lícita la pregunta que la Ciencia Ficción viene haciéndose hace más de dos décadas: ¿puede la tecnología de Inteligencia Artificial llegar a sustituir el pensamiento humano? Alan Turing ya se preguntaba algo semejante en su conocido artículo "Computing machinery and intelligence", de 1950. Dicho artículo comienza así: *"Propongo considerar la pregunta: ¿Pueden las máquinas pensar? Esto debería comenzar con definiciones sobre el significado de 'máquinas' y 'pensar'"*[3]. Según hipotetiza Turing, una máquina puede considerarse inteligente si logra engañar a los humanos haciéndoles creer que ella también es humana, y para discernir esto, Turing propone un juego, que luego se conocería como el "Test de Turing"[4].

---

[3] *"I propose to consider the question, "Can machines think? This should begin with definitions of the meaning of the terms 'machine' and 'think'".*

[4] Por supuesto que el Test de Turing fue muchas veces refutado, tal vez por no comprender acabadamente que lo que plantea Turing con su test no es si en verdad hay inteligencia —o algo que pueda denominarse antropomórficamente así— en la máquina, sino si los usuarios de la máquina (llamada "inteligente") creen o no que esa máquina tiene la posibilidad de "pensar" como lo hacen los humanos.

Mientras seguimos teniendo este interrogante solo en modo de posibilidad —que al día de hoy suponemos remota, pero ignoramos si mañana estará a la vuelta de la esquina— hay otro cuestionamiento urgente que debemos hacernos, y que nos coloca ante una posible frontera: debemos decidir como humanos qué lugar le daremos a la IA y si estamos dispuestos a delegar de forma consciente en esta tecnología emergente muchas de las actividades de nuestra vida cotidiana, creando de esta forma un nuevo territorio de posible convivencia entre inteligencias sintéticas, inteligencias artificiales, IA específicas[5], IA generalizadas[6] e inteligencias biológicas[7].

---

[5] La inteligencia artificial específica (o IA débil) se denomina de esta forma porque es capaz de realizar tareas de manera eficaz, incluso tal vez mejor que los seres humanos, pero esas tareas están acotadas, generalmente a un solo tipo de problema.

[6] La Inteligencia Artificial General (IAG) o "fuerte", es una forma teórica de IA que se caracterizaría por tener autoconsciencia, y posibilidad de comprender, aprender y realizar tareas intelectuales de forma muy semejante al cerebro humano. Por ejemplo: si una inteligencia artificial puede jugar al ajedrez, hasta el día de hoy, esa inteligencia artificial no podrá también jugar a las damas. Porque en la actualidad existe la llamada IA débil pero solo existe la IA fuerte como postulado teórico.

[7] Curiosamente, fue un filósofo del lenguaje, John Searle, quien distinguió por primera vez entre IA débil e IA fuerte. Según su

Así como Borges pone en jaque el concepto de "realidad" al mostrar cómo, en una suerte de cajas chinas, las realidades se solapan y se contienen unas a otras, se multiplican como en un laberinto de espejos y de esta forma se pierden las fronteras entre ellas, y entre ellas y la "realidad-real" que pierde su espesura e incluso su importancia, así también ocurre con las nuevas realidades que nos proponen las IA: nuestra percepción, creencias, imaginación y subjetividades humanas, de pronto se ven cuestionadas y conmocionadas. No hay respuestas definitivas sino

> Debemos decidir como humanos qué lugar le daremos a la IA y si estamos dispuestos a delegar de forma consciente en esta tecnología emergente muchas de las actividades de nuestra vida cotidiana, creando de esta forma un nuevo territorio de posible convivencia.

propuesta, solo el cerebro humano tiene IA fuerte, ya que es capaz de pensar de forma general. En sus palabras, las máquinas tienen "sintaxis", pero la mente humana tiene "sintaxis" y "semántica", lo que según él nunca tendrán las máquinas. (Searle, J. (1985). *Mentes, cerebros, ciencia*. Cátedra).

una invitación a la reflexión sobre la propia existencia y al desafío de los límites establecidos.

¿Es la Inteligencia Artificial la nueva *ficción* del siglo XXI? La palabra "ficción", en español, ha tomado la acepción de "invención" o "cosa fingida"[8], sin embargo, su origen etimológico proviene del latín *fingo-finxi-fictum*, cuyo primer significado es: "formar", "dar forma", "hacer", "modelar"[9], y por eso la pregunta es pertinente: ¿Qué realidad está *modelando* la IA enfrentándonos a la complejidad del conocimiento y a la búsqueda de sentido? ¿Qué realidad está *dando forma* a las entidades digitales sintéticas —que por el momento operan con "ecos" del conocimiento de los seres humanos, pero quizás con el tiempo operen con ficciones novedosas de una IA que alcance la

> ¿Puede la tecnología de Inteligencia Artificial llegar a sustituir el pensamiento humano?

[8] RAE.
[9] Diccionario VOX. Latín-Español Español-Latín. (1979). Bibliograf.

singularidad[10] o superinteligencia— transformando radicalmente la manera en que vivimos?

Hoy en día podemos plantear que la principal *ficción* en el desarrollo de la inteligencia artificial, se centra en obtener —a través de la programación evolutiva en integración con las redes neuronales

[10] El concepto de "singularidad" fue popularizado por Raymond Kurzweil en su libro de 2005 *La singularidad está cerca*, pero ya había sido planteado con anterioridad por el escritor de ciencia ficción Vernor Vinge (1993), quien auguraba el fin de la era humana, y mucho antes por el físico John von Neumann (1957). En 1965, el informático Irving Good afirmó que habría un evento al que denominó "explosión de la Inteligencia artificial", que consistiría en la última invención que el ser humano necesitara hacer, una "máquina ultrainteligente" que, además, debería indicarle al ser humano cómo mantenerla bajo control. La idea, entonces, no es nueva, pero fue modelándose a lo largo del tiempo y adquiriendo su significado actual, el cual hace referencia a la instancia de progreso (también llamada "tecnoptimismo"), cuando las tecnologías equiparen y superen a la inteligencia humana. Para Kurzweil, específicamente, las diferencias entre maquinas e inteligencia humana serán borradas a fines del siglo XXI. Esta "singularidad tecnológica" se implantará como consecuencia de lo que el autor llama "ley de rendimientos acelerados", por medio de la cual el mundo tal como lo conocemos será totalmente modificado de la mano de la nanotecnología, la genética y la robótica. En sus predicciones, Kurzweil sostiene que el ser humano podrá evitar de esta forma las limitaciones propias de la condición humana. La idea de "singularidad", además, contempla la posibilidad de que las propias IA se perfeccionen a sí mismas, en una espiral de evolución totalmente imparable. Hacia el 2029, predice Kurzweil, la IA, incluso, será capaz de pasar el Test de Turing.

digitales— algoritmos generativos basados en la construcción de sistemas cognitivos artificiales, fundamentados en el aprendizaje y el desarrollo de sus procesos con base al cerebro humano. Es decir, lo que se busca es lograr desarrollar y evolucionar hacia un cerebro artificial, cuya red neuronal se equipare al crecimiento y la maduración del cerebro humano.

Por ahora, debemos decir que solo permanece en el terreno de la ficción como simulación, e incluso como invención fantástico-literaria. Sin embargo, no podemos descartar que en el mediano o en el largo plazo —difícil, en virtud de la velocidad con que se nos presentan los cambios—lo aparentemente imposible pueda lograrse como una realidad una vez que la tecnología entre en juego.

Referencias bibliográficas

A. M. Turing (1950) Computing Machinery and Intelligence. *Mind* 49: 433-460. Recuperado de: https://phil415.pbworks.com/f/TuringComputing.pdf

González, R. (2007). "El Test de Turing: dos mitos, un dogma". *Revista de filosofía*, 63, 37-53. Recuperado de: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-43602007000100003

Good, I. J. (1965). I. (1965), "Speculations Concerning the First Ultraintelligent Machine". Recuperado de: https://web.archive.org/web/20010527181244/http://www.aeiveos.com/~bradbury/Authors/Computing/Good-IJ/SCtFUM.html

Kurzweil, R. (2012). *La singularidad está cerca*. Lola Books.

Searle, J. (1985). *Mentes, cerebros, ciencia*. Cátedra.

Vinge, V. (1993), "The Coming Technological Singularity", *Vision-21: Interdisciplinary Science & Engineering in the Era of CyberSpace, proceedings of a Symposium held at NASA Lewis Research Center*. Recuperado de: https://edoras.sdsu.edu/~vinge/misc/singularity.html

Capítulo 2

El destino algorítmico
Trascendiendo los límites de la cuarta dimensión

"En la era de la algoritmización total,
la realidad se pliega a las lógicas de la máquina,
erosionando las fronteras entre lo físico y lo virtual".
Eric Sadin

En la compleja y laberíntica red de la sociedad contemporánea, la fusión entre la educación y la inteligencia artificial (IA) ha surgido como un punto focal de innovación y transformación, dado el desarrollo exponencial de este tipo de tecnología. En este capítulo, exploraremos conceptos que desafían las nociones tradicionales acerca de la educación y que suscitan una reflexión profunda sobre el porvenir de la forma en que aprendemos y enseñamos en un mundo dominado por

algoritmos y entidades pseudo "pensantes" digitales y basado en el conectivismo[11].

Algoritmos del destino: la influencia de los algoritmos en nuestro devenir

La premisa de que los algoritmos modelan nuestro destino abre un debate existencial sobre el rol de la programación en la conformación de nuestras vidas. Hablar de "programación", acaso, es un

[11] El conectivismo es una etapa posterior a las tradicionales teorías sobre el aprendizaje, el conductismo, el cognitivismo y el constructivismo. Es posterior no solamente porque surge temporalmente después de ellas, sino porque en cierta manera las supera, al tomar de aquellas lo mejor e incorporar lo que la era digital a nivel global ofrece: la posibilidad de conexión, interconexión y globalización del saber a través de los dispositivos tecnológicos. El aprendizaje ya ha dejado de ser un fenómeno individual en tanto y en cuanto las nuevas herramientas permiten el desarrollo de nuevas habilidades y de nuevas interacciones entre quienes aprenden. El conectivismo es definido por Siemens (2007) como una nueva teoría del aprendizaje para la era digital. Así como la revolución industrial generó una teoría de enseñanza y aprendizaje acorde a las nuevas realidades que se estaban viviendo y a las nuevas exigencias que el mercado laboral imponía, la revolución digital también fija su impronta, al incorporar nuevas herramientas, nuevas concepciones antropológicas, nuevas necesidades laborales, nuevas teorías acerca de qué es, efectivamente, enseñar y aprender.

eufemismo para evitar pensar, ciertamente, que detrás de la "programación" existen personas y detrás de las personas intereses. Por eso, en un contexto donde la IA y los algoritmos son omnipresentes, resulta crucial indagar quién controla estos algoritmos y, como consecuencia de "quién/quiénes", qué valores están inscriptos en su código.

Tanto Foucault (1998) como Bourdieu (2001) explican las relaciones de poder, generalmente opacas, que rigen todas las acciones e interacciones sociales. El primero, dirá que la fuerza del poder radica en que "logra esconder sus mecanismos" (p. 93), y en este mismo sentido, Bourdieu dirá que las ciencias sociales son la "ciencia de lo oculto", es decir que su tarea crítica es poner de manifiesto aquello que está detrás de lo que se ve. Los conceptos aquí traídos a cuenta —de manera muy esquemática y

> ¿Hacia dónde va la tecnología y cómo afecta y afectará para bien y para mal la vida humana y todo lo existente? Detrás de los algoritmos hay poderes, y saberlo es imprescindible para el debate ético que sobre ellos se impone.

poco justa con la amplitud y profundidad con que los autores los desarrollan— nos sirven para reflexionar sobre el progreso de la tecnología, hacia dónde va y cómo afecta y afectará para bien y para mal la vida humana y todo lo existente, pero, sobre todo, nos sirven para advertir lo que dijimos más arriba: detrás de los algoritmos hay poderes, y saberlo es imprescindible para el debate ético que sobre ellos se impone.

Más allá de la eficiencia y la conveniencia que sin lugar a dudas la IA pone a disposición de todas las actividades humanas y, específicamente de la educación, la segunda pregunta que deberíamos hacernos es, ¿qué cosmovisiones perpetúan estos algoritmos? ¿Y cómo podemos garantizar que estas herramientas reflejen la diversidad y equidad que buscamos en la sociedad? Si, como dijimos, detrás de los algoritmos hay personas y entre las personas circula —opaco y oculto— el poder, entonces la siguiente cuestión a debatir será que, evidentemente, toda realidad conlleva sesgos ideológicos, culturales, contextuales en general, que condicionan la mirada sobre el mundo. Entonces: ¿qué cosmovisión, qué ideología, qué aspectos culturales, que

convicciones morales, qué paradigmas científicos o éticos, etc. existen detrás de esos algoritmos y, por tanto, los transmiten, los estratifican, los imponen? ¿Qué sucede, por ejemplo, si las manos invisibles detrás de esos algoritmos carecen de perspectiva de género? ¿Sería posible pensar en una IA que discrimine a las mujeres? ¿Y qué si las manos invisibles detrás de la IA son xenófobas, homófobas, aporofóbicas, por citar solo algunas posibilidades? ¿Obtendremos algoritmos que circulen por esos andariveles?

Los algoritmos, invisibles y omnipresentes, actúan como fuerzas determinantes que guían nuestras decisiones, preferencias y, en última instancia, nuestras vidas. Desde las recomendaciones de contenido en plataformas de *streaming* hasta las decisiones críticas en sectores como la salud, la economía y la justicia, los algoritmos están imbricados en el tejido de la sociedad moderna. Esta ubicuidad plantea preguntas profundas sobre la autonomía humana y la posibilidad de un destino predeterminado por código.

La Sociología y la Psicología social han realizado en los últimos cien años muchos experimentos sociales

para indagar hasta qué punto la conducta humana puede ser influida e incluso transformada por una fuerza exógena, es decir, cuán permeables somos los seres humanos a presiones externas, incluso si fueran en contra de nuestras propias convicciones.

En la década del ’60 del siglo XX, Stanley Milgram probó que la obediencia a una voz de mando podía ser superior a la fuerza interna para negarse a causar dolor a otros: los involucrados en la investigación no se negaban —aun cuando mostraban algunos reparos— a aplicarles descargas eléctricas crecientes a quienes oficiaban de “alumnos”.

En 1971, Philip Zimbardo ideó y llevó a cabo un experimento (“Cárcel de Stanford”) en el que se reclutó gente, a la que se le abonaba un estipendio, a quienes se asignó aleatoriamente el rol de guardias y el rol de prisioneros. Lo que se pudo observar, en este experimento que tuvo que cancelarse porque rápidamente se fue de las manos, es hasta qué punto los roles asignados condicionan las conductas de las personas y, además, cuánta influencia puede atribuirse al ambiente en las decisiones y comportamientos.

La idea de estos ejemplos, que podrían multiplicarse, aunque no es este el lugar para hacerlo, es mostrar que existen numerosos factores que son capaces de moldear la conducta de un ser humano en sociedad. Las diversas teorías sobre la propaganda y los medios en general dan cuenta de eso mismo: la teoría de Laswell de la aguja hipodérmica, explica cómo se puede manipular (gustos, preferencias, ideas, etc.) a través de pequeñas "inoculaciones" que el receptor ni siquiera nota, Harold Innis, por su parte, alerta sobre la estratificación que pueden provocar los medios masivos (y escribe cuando todavía la IA y la explosión en las comunicaciones estaba solo en los libros de Ciencia Ficción), Marshall McLuhan, hizo famosa su idea del mundo como una *aldea global,* Guy Debord y Jean Baudrillard describen la acción de las tecnologías en la conformación de una "sociedad del espectáculo" y de una sociedad como "simulacro".

En fin, desde mucho antes que apareciera en escena la IA, pensadores de muchas disciplinas se interrogan acerca de los contornos éticos y metodológicos de la aplicación de las tecnologías, tema

que en la actualidad está siendo intensamente debatido por filósofos, tecnólogos, sociólogos, politólogos y cientistas sociales en general.

Sin perder esto de vista, es más, teniendo estas previsiones en el horizonte de nuestras elucubraciones sobre la IA y sus relaciones con la educación, diremos, entonces, que la educación debe equipar a las personas no solo con habilidades para entender y desarrollar estos algoritmos, sino también con una conciencia crítica para cuestionar y desafiar las implicaciones éticas y sociales de su uso.

El despertar de las máquinas pensantes: algunos cuestionamientos

La metáfora de las máquinas adquiriendo conciencia plantea interrogantes fundamentales sobre la naturaleza de la inteligencia y la cognición. ¿Qué implica verdaderamente que una máquina "piense"? ¿Es meramente una simulación de procesos mentales o puede existir una auténtica conciencia artificial? Esta

pregunta no solo desafía nuestras concepciones sobre la inteligencia, sino también nuestra relación con las máquinas y la noción misma de humanidad.

El problema de las máquinas pensantes o de las máquinas con conciencia es un asunto muy tratado por la literatura y por el cine, incluso antes de que se produjera la revolución tecnológica a partir de las primeras computadoras. Los robots de apariencia humana, humanoides o androides, están en el imaginario artístico desde la pionera película de culto "Metropolis", de Friz Lang, de 1927.

Ahora bien, a pesar de ser un tema intuido, sospechado, anunciado por el arte muy antes de que aparezca, y ahora mismo muy debatido y estudiado, sin embargo, cabría preguntarse: ¿qué estamos diciendo, específicamente, cuando hablamos de una máquina con conciencia? En efecto, hay una dificultad clara al tratar de definir qué es, en realidad, "conciencia". Si atendemos a las definiciones de diccionario, el concepto va desde un aspecto moral y ético, es decir, algo así como el conocimiento del bien y el mal, hasta el hecho de percibir lo otro (algo así como "apercibirse de algo"), pasando por

un conocimiento profundo de lo que sea, o simplemente indica un estado de vigilia, frente a un estado de "adormecimiento". Como se ve, "conciencia" es un término polisémico que abarca un arco de significación amplio, ambiguo y hasta contradictorio. Para la ciencia, mientras tanto, el panorama no es más claro: la conciencia se debe a un complejo entramado neuronal, pero no hay todavía certeza de cómo se produce y por qué. Podría decirse que la conciencia es la percepción de todo lo que no es el yo, lo que está fuera de uno mismo, pero también autopercibirse en sentimientos, gustos, emociones, necesidades, etc., y esa percepción y autopercepción tienen como sujeto/objeto a un yo que permanece estable. ¿Qué es la conciencia, entonces? Algo bastante difuso, todavía un misterio.

> La educación debe equipar a las personas no solo con habilidades para entender y desarrollar estos algoritmos, sino también con una conciencia crítica para cuestionar y desafiar las implicaciones éticas y sociales de su uso.

Si queremos abordar el concepto de "conciencia" desde un aspecto filosófico, la cuestión no resulta más

sencilla. En la entrada correspondiente del Diccionario de Filosofía de Ferrater Mora (1958, p. 323) se habla de dos sentidos fundamentales: *percatación o conocimiento de algo* o discriminación entre el bien y el mal. En cuanto al primer sentido, según el autor puede interpretarse en tres planos: psicológico, epistemológico y metafísico. El aspecto psicológico remite a la conciencia de sí, de la propia existencia, el aspecto epistemológico se relaciona con el modo en que se conoce lo que no es uno mismo, los mecanismos mentales que intervienen, y la interpretación metafísica de la conciencia apunta a algo que está antes de lo psicológico o epistemológico: es el yo mismo.

Los filósofos no se ponen de acuerdo acerca de si la conciencia es una función, una facultad, o una actividad psíquica.

Como vemos, definir a qué nos referimos cuando hablamos de "máquinas conscientes" o "máquinas con conciencia", no es una tarea sencilla.

Si, por otra parte, hablamos de "máquinas pensantes", tampoco el panorama está resuelto: ¿qué es pensar? La Rae dice que es formar y combinar ideas y examinar mentalmente

algo. El mismo diccionario de Ferrater Mora dice que "pensar" es una actividad interna que tiende a aprehender objetos (no solo objetos tangibles). A excepción de que es una actividad interna y de que se relaciona con ideas, todo respecto del pensar es materia de investigación y análisis: cómo se produce, por qué se produce, cuál es su sustancia, cuáles sus elementos, etc.

Nuevamente, *pensar* en "máquinas pensantes" encuentra los mismos escollos epistemológicos al momento de definir qué son realmente. Quizás esta nebulosa —necesario es decir que los adelantos científicos respecto del cerebro humano son cualitativamente y cuantitativamente enormes en esta época respecto de otras— hace imperativo que planteemos nuestras dudas y previsiones que creemos necesarias respecto de la revolución que ya se ha desatado a la espera de una IA fuerte.

En este contexto, la educación desempeña un papel esencial al preparar a las próximas generaciones para interactuar con estas entidades emergentes —que ya están entre nosotros y nosotras y será imposible eludirlas— pero para que esas interacciones se realicen

de forma ética y significativa. Por ello, es oportuno, en este momento, citar a Ortega y Gasset (1962, p.32), para quien "la vida humana constituye el nervio principal de la metafísica existencial". Según él, ser humano es una "realidad radical"; radical porque todos los demás tipos de realidad —tanto física como espiritual y hoy podríamos agregar "virtual–digital"— dependen de la existencia de las personas.

La emergencia de máquinas pensantes requiere, por tanto, una revisión profunda de los currículos educativos. Los y las estudiantes deben ser expuestos no solo a la programación y el desarrollo tecnológico, sino también a la filosofía, la ética, la antropología y la sociología. Entender qué significa la "conciencia" en un contexto artificial podría revolucionar nuestras concepciones de la mente y la identidad, y para ello, la educación debe facilitar una comprensión holística que abarque tanto los avances científicos como las implicaciones humanas de vivir junto a entidades con capacidades cognitivas avanzadas.

Ficciones y perspectivas de la IA: trascendiendo límites

La aparición del metaverso en el horizonte de la realidad tal como hasta ahora la conocíamos es la emergencia de lo que está "junto a" o "más allá" (μετά) del universo físico. En todo caso, es otro *universo*. Una verdad obvia sería plantear que, si no hemos podido conocer hasta ahora el universo todo, el metaverso implica, incluso, un desafío más grande. Sin embargo, ya está allí, μετά: lo veamos o no, lo queramos ver o no, seamos conscientes o no de sus implicancias en nuestra vida diaria.

Existe ya una expansión de la realidad hacia dimensiones digitales. Por eso, la concepción del metaverso como una nueva esfera de existencia plantea desafíos emocionantes y desconcertantes para la educación, en cuya convergencia debemos considerar: ¿cómo enseñamos en un espacio que desafía las leyes físicas y las normas sociales convencionales? ¿Qué oportunidades y

peligros aguardan en esta frontera entre lo real y lo virtual? Estos interrogantes demandan una reimaginación radical de los métodos educativos y una mayor atención a la ética y la inclusión en el diseño de entornos virtuales de aprendizaje.

El metaverso, una realidad digital inmersiva, abre un sinfín de posibilidades para la educación. Espacios de aprendizaje que trascienden las limitaciones físicas pueden ofrecer experiencias personalizadas y envolventes. Sin embargo, este nuevo terreno también está plagado de riesgos, desde la desconexión de la realidad física hasta la exacerbación de desigualdades preexistentes. La educación en el contexto del metaverso y la integración de la inteligencia artificial dentro del mismo debe ser inclusiva y accesible, asegurando que todos los estudiantes, independientemente de su origen socioeconómico, puedan beneficiarse de estas innovaciones.

> La educación desempeña un papel esencial al preparar a las próximas generaciones para interactuar con estas entidades emergentes — que ya están entre nosotros y nosotras y será imposible eludirlas — pero para que esas interacciones se realicen de forma ética y significativa.

El planteo sobre la igualdad de oportunidades, sobre la inclusión, sobre la posibilidad de sesgos ideológicos o de nuevas desigualdades 2.0, 3.0 o 4.0, en una suerte de nueva "selección natural" o "supervivencia del más —o mejor— conectado", se da en un contexto paralelo a la discusión, en ámbitos educativos, sobre la pertinencia o no de la enseñanza de las ciencias sociales, interpretadas por algunos como "poco productivas" en un mundo tendiente a ser dominado por las tecnologías.

Trascender los límites, cruzar los portales, ya no parece ser solo una opción. Hacerlo desde un pensamiento crítico —ese provisto justamente por las ciencias sociales despreciadas por los tecnócratas— tampoco es solo una opción: se presenta más bien como un imperativo tendiente a minimizar riesgos y efectos indeseados.

Desafíos y oportunidades: navegando en la Cuarta Dimensión

Pero este no es un libro sobre los riesgos (que no eludimos, sino que ponemos sobre la mesa en toda su crudeza) sino sobre las oportunidades. Porque el surgimiento de las máquinas pensantes plantea desafíos multifacéticos para la educación. Preparar a los individuos para esta nueva era exige un enfoque integral que abarque no solo habilidades técnicas, sino también competencias socioemocionales y éticas. La autonomía, la creatividad y la adaptabilidad se tornan cruciales en un mundo donde la automatización redefine los límites del trabajo humano. Además, la educación debe ser un faro de esperanza en medio de la oscuridad de la desigualdad digital, asegurando que todos tengan acceso equitativo a las oportunidades que ofrece la era de la IA.

La intersección entre la educación, el metaverso y la IA no solo representa un desafío, sino también una oportunidad sin precedentes para redefinir el paradigma educativo y preparar a las generaciones

futuras para un mundo en constante evolución. Al abrazar la incertidumbre y la complejidad de este viaje, podemos trazar un camino hacia un futuro más inclusivo, equitativo y lleno de posibilidades. En este camino, la educación debe ser un espacio de innovación constante, donde se cultiven tanto las habilidades técnicas necesarias para manejar la tecnología como las cualidades humanas que nos permiten convivir y prosperar en una sociedad cada vez más digitalizada.

El rol del educador, por supuesto, debe transformarse, pasando de ser un mero transmisor de conocimientos a un guía que ayuda a los estudiantes a navegar por un mundo lleno de información y posibilidades.

Este nuevo horizonte educativo, enriquecido y a la vez desafiado por la inteligencia artificial, nos invita a reflexionar y actuar con visión y responsabilidad, asegurando que el futuro de la educación sea tan brillante y justo como las tecnologías que lo modelan.

Referencias bibliográficas

Bourdieu, P. (2001). *Poder, derecho y clases sociales*. Desclée de Brouwer.

Debord, G. (1967). *La sociedad del espectáculo*. Buchet-Chastel.

De la Fuente, R. El estudio de la conciencia: estado actual. *Salud Mental*, vol. 25, núm. 5, octubre, 2002, pp. 1-9. Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz. Distrito Federal, México.

Ferrater Mora, J. (1958). *Diccionario de filosofía*. Sudamericana.

Foucault, M. (1998). *Historia de la sexualidad I. La voluntad de saber*. Siglo XXI.

Foucault, M. (2009). *El orden del discurso*. Tusquets.

Foucault, M. (2014) *Las redes del poder*. Ed. Prometeo.

Foucault, M. (2014) *Las redes del poder*. Ed. Prometeo.

Gutiérrez, L. (2012). Conectivismo como teoría de aprendizaje: conceptos, ideas, y posibles limitaciones. *Revista Educación y Tecnología*, N° 1, año 2012.

Hoyuelos, M. (2020). Conciencia artificial en la ciencia y en la ficción. *Núcleos #7*. Diciembre de 2020

Lasswell, H. (1927). *Propaganda Technique in the World War*. A. A. Knopf.

McLuhan, M. (1972). *La Galaxia Gutenberg*. Editorial Aguilar.

Siemens, G. (2007). *Conectivismo: Una teoría de aprendizaje para la era digital.* Humanas virtual.

Ortega y Gasset, J. (1962). Historia como sistema. En *Obras completas*. Revista de Occidente.

Capítulo 3

Los Golem Digitales

Prospectiva y perspectivas de la IA

—Golem, no estás formado del todo,
pero ahora te voy a terminar.
Sabe que fuiste creado para poco tiempo
y con un propósito.
No intentes nunca salirte de esa senda.
Harás lo que yo te mande.
—Como has sido creado con un único fin, se te ha dado
un cerebro diferente del de los hombres.
Sin embargo, nunca se sabe cómo funciona un cerebro.
Al llegar a casa, el rabino recitó la oración de la noche
y se acostó.
Por primera vez en muchos años no se pudo dormir.
Se le había concedido un gran poder desde el cielo, y
tenía miedo de no haberlo merecido.
(…)
Le parecía al rabino que los ojos del gólem
preguntaban: «¿Quién soy yo?
¿Por qué estoy aquí? ¿Cuál es el secreto de mi ser?».
(…)
—¿Has oído lo que te he mandado hacer?
Y el gólem dijo: —Sí.
—¿Lo harás? —preguntó rabí Leib.
Y el gólem repuso:
—No.
—¿Por qué no? —preguntó, asombrado, rabí Leib.
Pareció como si el gólem reflexionara por unos
instantes,
y luego dijo:

—Gólem no saber.
(…)
Como rabí Leib había prometido al emperador
poner fin a la existencia del
gólem, le dijo:
—Baja la cabeza.
Tenía la intención de borrar el Santo Nombre de la frente del gólem, ya para siempre. Pero, en vez de bajar la cabeza,
el gólem dijo:
—No.
Rabí Leib comprendió claramente que había perdido para siempre la autoridad que tenía sobre el gólem.
El rabino se afligió. De nada serviría discutir con un gólem que no tenía entendimiento.
Rabí Leib había cometido un error
que ya no podía corregir.

Golem. El coloso de barro. Isaac Bashevis Singer[12]

El origen de la leyenda del Golem, como todo mito, leyenda o historia fantástica y folklórica primero sostenida por tradición oral, es de muy difícil rastreo, porque entronca, como ya nos explicaran los Formalistas Rusos, especialmente a través de los trabajos de Vladimir Propp (1998), con estructuras de pensamiento —llamadas "arquetipos" por Karl Gustav Jung (1974)—

[12] (1901-1993). Escritor judío de origen polaco, Premio Nobel de Literatura en 1978.

que se repiten en diferentes pueblos y civilizaciones aunque no se relacionen entre sí ni espacialmente, ni culturalmente ni temporalmente. Son como formas de pensamiento arcaicas que se repiten, modos en que los seres humanos de todos los tiempos intentamos explicar los arcanos de la realidad.

> ¿Podemos los seres humanos convivir con los Golem digitales teniendo consciencia de los lugares que les estamos permitiendo manejar, controlar y, por qué no, dominar, siendo que estas nuevas entidades carecen de inteligencia propia y consciencia?

En cuanto al Golem, si bien con ese nombre corresponde a la narrativa judía medieval, también podemos encontrar la leyenda (con particularidades) en otras civilizaciones. Su versión del siglo XVI es atribuida al Maharal de Praga, Yehuda Loew b. Bezalel, un rabino experto en esoterismo, alquimia y otras disciplinas quien, ante la amenaza de persecución y matanza de los judíos con acusaciones falsas de haber matado y desaparecido a un niño cristiano —eran tiempos convulsionados de guerras religiosas—, creó una criatura de arcilla y barro para que sirviera y defendiera a la co-

munidad judía. La criatura cobró vida con las artes alquímicas del Maharal, poniéndole en su boca un papel con el nombre secreto de Dios (algunos dicen que era la palabra "Emeth" —verdad, veracidad, característica normalmente atribuida al Dios judío, junto con "hesed", gracia, amor incondicional—). El Golem se activaba y se desactivaba poniéndole o sacándole ese papel con el nombre sagrado de la "boca". La idea, detrás de esta leyenda, es la posibilidad de crear, de "jugar a ser Dios", que no es nueva ni exclusiva de los pueblos que se asumen religiosos. Es una aspiración arquetípicamente humana que está narrada en muchas leyendas y literaturas de corte fantástico de larga data. En la leyenda judía, cierto día el "creador" no sacó el "papel" de la "boca" del Golem, y este se descontroló, ocasionando toda suerte de desmanes, hasta que pudo dominarse mediante la extracción de la "palabra mágica". Mucho tiempo después, en el poema homónimo de Jorge Luis Borges, uno de los versos da la clave de por qué alguien se embarcaría en semejante creación, dice: *"Sediento de saber lo que Dios sabe"*. No importa, en este caso, la creencia en Dios, sino

en lo que eso simboliza: llegar a tener todo el poder y todo el saber. Eso sería "llegar a ser como Dios".

Uno de los finales atribuidos al Golem narra que el rabino quitó la palabra "emeth" y la cambió por "meth" —muerte— y así el coloso se desmoronó. Y otra conclusión que se conoce de la leyenda es que, en su descontrol, el Golem destruyó a su hacedor.

En el poema de Borges, en cambio, no aparece este descontrol y esta destrucción, pero sí hay en énfasis en mostrar al Golem como una criatura sin inteligencia propia, sin habla (con todo el simbolismo que representa el habla y el lenguaje), tosco, incapaz de pensar por sí mismo.

¿Cuáles son los elementos interesantes de destacar en esta leyenda que nos sirve como disparador? Pues bien, el Golem surge como una creación humana, creación proveniente de la *hybris*, es decir, de la desmesura, de la arrogancia del creador o hacedor. Esa criatura, no tiene "alma", no piensa, no tiene inteligencia propia y, en esa tosquedad, es capaz, incluso, de volverse en contra de su propio creador.

Entrando ya al tema que nos ocupa, podemos hacer un paralelo entre esta creación artificial del Golem —aspiración que encarna arquetípicamente a la de la humanidad de todos los tiempos—, y las implicaciones éticas que conlleva, con la creación de una Inteligencia Artificial, es decir, de desarrollos de entidades digitales inteligentes[13] que, como el Golem y con el transcurso del tiempo, se volverán poderosas y hasta cierto punto obedientes. Al crearlas, intentamos darles sentido justificándonos en la necesidad de avances, los cuales son muy buenos para el desarrollo de las sociedades y de las naciones.

Pero la pregunta que no podemos dejar de hacernos, tiene que ver con mucho de lo que nos relata la leyenda del Golem, esa criatura que carecía de inteligencia propia y de conciencia, pero que se volvía cada vez más poderosa: ¿podemos los seres humanos

---

[13] Entidades digitales inteligentes, son sistemas o programas informáticos que utilizan tecnologías como la inteligencia artificial y el aprendizaje automático para interactuar y comunicarse con los usuarios de manera autónoma. Estas entidades tienen la capacidad de comprender el lenguaje natural, procesar información, tomar decisiones y realizar tareas específicas sin intervención humana directa.

convivir con los Golem digitales teniendo consciencia de los lugares que les estamos permitiendo manejar, controlar y, por qué no, dominar, siendo que estas nuevas entidades carecen de estas dos habilidades fundamentales? ¿Somos capaces, como humanos, de mantener un control sobre estas entidades? ¿Contamos con mecanismos que nos permitan retomar el control en caso de una rebelión de estos Golem digitales? Y la pregunta inquietante que no podemos dejar de realizarnos: ¿es la Inteligencia Artificial —o puede llegar a ser— una suerte de Golem que escape de las manos creadoras con fines no previstos o en cambio tenemos todas las herramientas necesarias para subordinarla a los intereses humanos y someterla a parámetros que no puedan autárquicamente trascender?

La Inteligencia Artificial (IA) —es decir estas nuevas entidades digitales inteligentes— ha experimentado un crecimiento exponencial en las últimas décadas, convirtiéndose en una de las tecnologías emergentes más prometedoras, pero a la vez más disruptivas y conflictivas de nuestro tiempo. A medida que la IA continúa evolucionando, es importante

analizar su prospectiva y las perspectivas asociadas con su desarrollo, así como su impacto en la vida de los seres humanos. Este libro se propone como una apertura a la reflexión sobre estos puntos, y como una mirada crítica que —no por crítica— deja de ser esperanzadora y positiva respecto de las posibilidades que el futuro (y el presente) nos ofrece.

La IA se ha integrado en múltiples aspectos de nuestras vidas, hecho que configura una bisagra tecno social para un futuro-presente, que ya hemos comenzado a vivir.

La prospectiva de la Inteligencia Artificial se presenta como un campo amplio y en constante expansión. Las investigaciones —y sus consecuentes avances en IA— han llevado a la creación de sistemas cada vez más sofisticados, capaces de realizar tareas que antes solo los seres humanos podían realizar. Desde la conducción autónoma hasta la traducción automática o la atención médica remota, asistida por Inteligencia Artificial, entre tantas otras. Tanto es así que la IA se ha integrado en múltiples aspectos de nuestras vidas,

hecho que configura una bisagra tecno social para un futuro-presente, que ya hemos comenzado a vivir.

No obstante, el desarrollo de la IA también plantea desafíos y perspectivas que deben ser considerados. Por un lado, existe el temor de que la IA pueda reemplazar a los seres humanos en el ámbito laboral, lo que podría resultar en la pérdida de empleos y la desigualdad económica. Sin embargo, también se argumenta que la IA tiene el potencial de mejorar la eficiencia y liberar a los seres humanos de tareas monótonas, peligrosas y repetitivas, permitiéndoles enfocarse en actividades más creativas y significativas. Al respecto podríamos decir que ambas cosas son reales, o potencialmente reales: desde la Revolución Industrial estamos asistiendo a un fenómeno de sustitución de mano de obra humana a causa de la tecnificación. Tareas que antes requerían de varias personas para su proceso, ahora demandan apenas de una o de ninguna. Supongamos, por ejemplo, el empaquetado de productos: una línea de producción montada en una o varias máquinas concatenadas pueden cumplir con la tarea igual o mejor que muchas

manos humanas. Es cierto y es una realidad que no puede pararse. La industria se ha reconvertido muchas veces a lo largo de estos años, y los trabajos se han tenido que hiperespecializar para dar respuesta a los nuevos escenarios del mercado laboral. El Foro Económico Mundial estima que en los próximos años la automatización se cobrará la baja de más de 85 millones de empleos, en lo que ya se denomina "Cuarta revolución industrial"[14]. El resultado es, por cierto, paradójico: hay una mejora en la productividad, aparecen nuevos bienes y servicios y, por tanto, se modifica la estructura económica global, sin embargo, y a la vez, se produce como consecuencia un desplazamiento humano del mercado laboral. Para Kelly (2016) todos los cambios que ofrece esta nueva revolución industrial se relacionan, directa o indirectamente, con la tecnología:

-Progreso y difusión de la IA.

[14] https://es.weforum.org/press/2020/10/la-recesion-y-la-automatizacion-cambian-nuestro-futuro-laboral-pero-hay-trabajos-que-se-avecinan-afirma-un-informe/

-Transmisión de conocimiento en tiempo real.

-Evolución de los productos, bienes y servicios y acortamiento del plazo de obsolescencia.

-Producción personalizada (posible por las nuevas tecnologías) que tienden a suplantar la producción en masa.

-Colaboración permanente entre usuarios y creación de *comunidades colaborativas* que, a su vez, coadyuvarán al crecimiento y desarrollo de tecnologías.

-Virtualización de la vida.

-Conexión cada vez más desarrollada persona-máquina y máquina-persona.

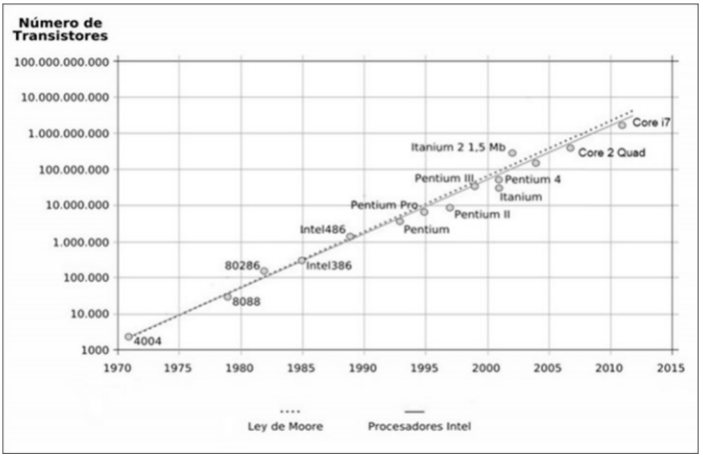
El progreso es tan vertiginoso que ya Gordon Moore había predicho, en 1965, que la tasa de transistores por centímetro cuadrado en un sistema integrado se duplicaría año a año, y luego corrigió esa hipótesis a año y medio, en un crecimiento sostenido de la nanotecnología que ha venido comprobándose a lo largo del tiempo. Este progreso constante y sostenido ha dado en llamarse, desde entonces, "Ley de Moore"[15]. Por

[15] Cuantos más componentes haya en un dispositivo más pequeño, más prestaciones brindará y a menor precio. Veamos este gráfico

lo tanto, ya no hay que esperar a un futuro lejano para que se produzcan altas tasas de desocupación, por un lado, y una modificación total en el mercado laboral, por el otro, cuyas demandas específicas tienden a la hiperespecialización, es decir, a la especialización acotada a un nicho particular del conocimiento.

La OCDE[16] y la CEPAL[17] son optimistas, sin embargo, en pronosticar que el desempleo solo

como ejemplo de la evolución de los procesadores Intel. (https://www.emibin.com/ley-de-moore-intel-10nm/)



[16] OCDE: Organización para la Cooperación y el Desarrollo Económicos. Organización internacional que colabora con los gobiernos para promover políticas que tiendan a la equidad, a la igualdad y al bienestar de las personas.

[17] CEPAL: Comisión Económica para América Latina dependiente de Naciones unidas, que funciona desde 1948. Su función es contribuir

descenderá 1% o 2% a causa de la tecnificación, y a la vez es auspiciosa en afirmar que se crearán nuevos empleos (hasta ahora poco conocidos o absolutamente desconocidos) a causa de las tecnologías emergentes[18]. El tema de IA-desempleo o IA-nuevos empleos es un tópico que está discutiéndose todavía y al que solo podremos darle un cierre conforme se desarrollen los acontecimientos. Sin embargo, es necesario afirmar en este punto que para que lo segundo suceda (IA-nuevos empleos), será necesario un cambio rotundo en la educación, en la política educativa y, sobre todo, en la accesibilidad de los estudiantes a los nuevos dispositivos, cuestión que, en Latinoamérica, no es un detalle menor a debatir.[19]

Ahora bien, unos párrafos más arriba hablábamos de que la posibilidad de trabajar menos

---

al desarrollo de América Latina y el Caribe, promocionando y coordinando las relaciones económicas entre los países latinoamericanos entre sí y con el mundo.

[18]https://www.argentina.gob.ar/noticias/nuevas-tecnologias-y-trabajos-del-futuro-la-urgencia-de-acortar-la-brecha-digital-para

[19] A este respecto es interesante el concepto de "Robotlución", aportado por el BIR (Banco interamericano de Desarrollo), que analiza el impacto de la automatización en la matriz productiva.

abría el juego al empleo de las horas libres en tareas de esparcimiento o en tareas más creativas. Esto es verdad, o quizás sea una media verdad en economías deprimidas como las que padecemos en América del Sur, en las que la tasa de desempleo (sin contar la producida por la IA) es a veces de dos dígitos, donde hay tantas personas bajo la línea de la pobreza y donde el problema mayor no es la tasa de crecimiento del país sino la redistribución de la riqueza. Todo cambio auspicioso debe ser analizado críticamente y contextualmente: no será el mismo efecto el que provocará el avance de la IA en Finlandia que en Argentina o Perú. Por otro lado, también la IA nos enfrenta a interrogantes éticos y sociales. El desarrollo de sistemas de IA autónomos plantea preguntas sobre la responsabilidad y la rendición de cuentas en caso de que ocurran errores o decisiones perjudiciales. También

> Es posible, entonces, pensar un escenario en el que aquellos que por su poder económico quiebren de alguna manera la idílica "democratización del conocimiento" que proponen las nuevas tecnologías, puedan también imprimirle un sesgo.

existe la preocupación de que la IA pueda perpetuar sesgos y discriminación, ya que los algoritmos aprenden de datos históricos que pueden reflejar prejuicios existentes en la sociedad: puesto que es un dato fáctico que la IA no se autogenera, sino que siempre parte de datos y conocimiento producidos por personas, es posible, entonces, pensar un escenario en el que aquellos que por su poder económico quiebren de alguna manera la idílica "democratización del conocimiento" que proponen las nuevas tecnologías, puedan también imprimirle un sesgo. No necesariamente sea "malo", "antidemocrático", "racista" o discriminador en general… Pero puede serlo. Existe como posibilidad. Por eso es crucial abordar estos desafíos para garantizar un desarrollo ético y responsable de la IA.

El impacto de la IA en la vida de los seres humanos es innegable. En el ámbito de la salud, por ejemplo, la IA ha demostrado su capacidad para diagnosticar enfermedades con mayor precisión y rapidez, lo que puede salvar vidas y mejorar la atención médica. Del mismo modo, en el campo de la educación,

la IA puede personalizar la enseñanza y adaptarla a las necesidades individuales de los estudiantes, brindando una experiencia de aprendizaje más efectiva y accesible.

Sin embargo, también es importante considerar los posibles efectos negativos. El uso masivo de la IA genera dudas sobre la preservación de la privacidad y la seguridad de los datos personales. Cualquier usuario de tecnologías puede comprobar a diario que recibe en sus dispositivos publicidad acorde a sus gustos y preferencias, lo que hace suponer, de mínima, que las búsquedas en internet no son inocuas y que, por tanto, hay algo de la privacidad que se pone en juego cada vez.

Por último, existe el riesgo de una dependencia excesiva y de la consecuente delegación de muchas de las decisiones que hoy tomamos libremente como humanos, convirtiéndolas en acciones decisorias a través de entidades digitales. Esto podría significar que nuestras decisiones y acciones —incluso por el desconocimiento sobre estos avances— puedan ser influidas o encaminadas —*ex profeso* o de alguna forma ingenua—. Entonces, cabe interrogarse: ¿nos

encontramos entregando el “dial del destino humano” a las máquinas?

Asimismo, y en prospectiva, es urgente preguntarnos de manera consciente si estamos ante escenarios en los que podríamos estar permitiendo a los cerebros artificiales, estos nuevos Golem digitales, tener el poder y el inquietante rol de gobernar a los seres humanos y las cosas.

Referencias bibliográficas

Arellano Espinar, A., Díez Catalan, L. y Neut Smith, R. (2019): "Revolución tecnológica y desarrollo socioeconómico: una agenda para el mercado de trabajo". *Un Nuevo Contrato Social en una Nueva Economía*, noviembre-diciembre 2019, N°911.

Argüello Verbanaz, S. (2019): "Los efectos de la automatización sobre el trabajo. Desempleo tecnológico, polarización del mercado laboral y políticas públicas". *BCN*, Nº SUP: 118.701.

Arntz, M., Gregory, Y. y Zierahn, U. (2020): "Digitalization and the future of work: macroeconomic consequences". In Zimmermann, K. (ed). *Handbook of Labor, Human Resources and Population Economics.* Forthcoming.

Bashevis Singer, I. (2011). *Golem, el coloso de barro*. Noguer ediciones

BID (2017): "Robotlución: El futuro del trabajo en la integración 4.0 de América Latina". *BID-INTAL*, Buenos Aires, República Argentina.

Calveira, M. Fraccia, E. (2020). Impacto de la cuarta revolución industrial en el empleo y en la distribución del ingreso. Recuperado de: https://ideas.repec.org/p/aep/anales/4346.html

Cheang Wong, J. (2005). Ley de Moore, nanotecnología y nanociencias: síntesis y modificación de nanopartículas mediante la implantación de iones. *Revista Digital Universitaria.* 10 de Julio 2005. Volumen 6 Número 7. ISSN: 1067-6079Recuperado de: http://www.revista.unam.mx/vol.6/num7/art65/int65.htm

Jung, C. G. (1999). *Obra completa*. Editorial Trotta S.A.

Kelly, K. (2016). *The inevitable: Understanding 12 technological forces that Will shape our future.* Viking.

Propp, W. (1998). *Morfología del cuento.* Akal.

Varela Moreno, M. (1995). La leyenda del Golem. Orígenes y Modernas derivaciones. *BIBLID* [0544-408X. (1995)44;61-79]

Capítulo 4

Horizontes de la Inteligencia
Desentrañando el significado detrás de la Inteligencia artificial y humana

Es la ciencia de hacer que las máquinas hagan cosas que requerirían inteligencia si las hubiera hecho un humano.

Marvin Minsky

¿No se te ha ocurrido pensar, Winston, que lo más tarde hacia el año 2050, ni un solo ser humano podrá entender una conversación como ésta que ahora sostenemos?

George Orwell, *1984*

Primeramente, tenemos que establecer qué es la inteligencia artificial y cómo se estructura en base a su desarrollo tecnológico. En el marco de las Ciencias de la Computación o Ciencias de la Informática, y de modo amplio, se entiende a la Inteligencia Artificial (IA) como un conjunto de saberes o una disciplina específica dentro del conglomerado de las ciencias computacionales.

Sin embargo, y específicamente, la IA hace referencia a un conjunto de capacidades —epistemológicas, cognosciti-

vas, intelectuales— reconocibles en sistemas informáticos y máquinas, cuyas combinaciones de algoritmos imitan la inteligencia humana para realizar una amplia gama de tareas, desde las más mecánicas y simples, hasta las más complejas. Asimismo, una característica distintiva de la IA es su potencialidad —casi sin frontera—, de evolucionar, mejorar e incluso abrirse a resultados no esperados (buenos o malos), a medida que aumenta la información que recopila y también a medida que la combina entre sí. En este punto es necesario remarcar que la IA es, siempre, una imitación, una emulación de la inteligencia humana, afirmación que, por sí sola, no elimina de raíz los planteos éticos que puedan realizarse, puesto que dichos planteos no están dirigidos específicamente a esta condición de "imitación" sino a las derivas que la IA pueda tener a futuro, y a la posibilidad —cierta o no— de que de alguna manera y en algún momento, logre independizarse de su modelo. Por el momento, digamos, la IA funciona a modo de "imitación", y en esto habría que diferenciarla de la "Inteligencia sintética" planteada por

algunos teóricos. En efecto, fue John Haugeland[20] en 1985 quien propuso este concepto, al ir un poco más allá de la simple "emulación" de la inteligencia humana. La idea, según Haugeland, es conseguir "máquinas que tengan mente", no solamente "lograr un simulador ingenioso"[21], porque, según su línea argumentativa, esa inteligencia lograda "artificialmente" es, de alguna manera, una inteligencia diferente, fruto de una "síntesis" entre la inteligencia humana y "algo más". No obstante, el propósito de la IA, en principio, sería mejorar prostéticamente —mediante dispositivos externos— las capacidades que naturalmente tienen los seres humanos. Y decimos "en principio", puesto que no es el objetivo de máxima suplantar la inteligencia humana, y tampoco la

[20] (1945-2010). Filósofo especializado en Filosofía de la mente, Fenomenología y Filosofía de la Ciencia,

[21] "*Tal vez la 'inteligencia artificial' se debería llamar 'inteligencia sintética' para que concuerde mejor con el lenguaje comercial. Así, los diamantes artificiales son falsas imitaciones, mientras que los diamantes sintéticos son diamantes auténticos, solo que manufacturados en lugar de desenterrados (podríamos comparar también el sabor artificial a maple con, digamos, la insulina sintética). No obstante, el nombre, la inteligencia artificial aspira a una inteligencia auténtica, no a una falsa imitación".* (Haugelman, J, 1985).

de Haugeland es la comprensión de la IA que más adeptos tiene.

> IA hace referencia a un conjunto de capacidades —epistemológicas, cognoscitivas, intelectuales— reconocibles en sistemas informáticos y máquinas, cuyas combinaciones de algoritmos imitan la inteligencia humana.

No hay una definición unívoca de lo que la IA es, pero las posturas mayoritarias podrían englobarse en la consideración de que la IA son sistemas que piensan y actúan como humanos, por un lado, o sistemas que piensan y actúan racionalmente, por otro. Entre ambos, el lector podrá advertir una sutil diferencia. El elemento "humano", además del raciocinio, supone otros aditamentos que tienen que ver con la común "condición humana", cuya imperfección y precariedad hacen posible, incluso, los errores. Como quiera que sea, la IA como disciplina científica se dedica al diseño de agentes inteligentes que puedan desarrollar actividades y tareas que requieren de inteligencia, incluso con la hipótesis de

máxima de lograr que la *performance* de esos agentes sea incluso mejor que la de los seres humanos.

El cerebro humano piensa, actúa, toma decisiones por medio de redes neuronales compuestas por más de cien billones de neuronas e infinitas interconexiones entre ellas. La IA imita estas redes neuronales con sistemas de *input* y *output* de datos en diferentes "capas" y sofisticados niveles de procesamiento. El *"Deep learning"* o *"Deep network"* hace referencia a las capas más profundas de estas redes neuronales, cada una de las cuales analiza y procesa datos con finalidades específicas. Los algoritmos en la IA tienen la capacidad de clasificar características e identificar patrones que luego serán útiles para los fines propuestos de esa IA, y, a la vez, esos algoritmos poseen la capacidad de aprender de los datos, hacer predicciones y tomar decisiones (*Machine learning).*

Llegados a este punto, y ya que hemos mostrado la relación que existe entre la IA y la inteligencia humana —ya sea como emulación o como síntesis y superación— deberíamos preguntarnos seriamente con qué concepto de "inteligencia" nos estamos manejando

o, más específicamente, qué tenemos como horizonte de sentido cuando decimos "Inteligencia artificial": ¿a qué inteligencia nos referimos? ¿qué inteligencia emula o supera la IA?

Por supuesto, somos plenamente conscientes de que no podemos agotar este planteo en unas pocas líneas en un libro cuyo tema no es el constructo "inteligencia" y todas sus implicancias psicológicas, psicométricas, pedagógicas, sociológicas, etc., sino la Inteligencia Artificial. No obstante, dedicaremos solo unas líneas para advertir al lector acerca de las múltiples aristas que tiene el problema y el complejo perspectivismo desde el cual se puede abordar la temática y sus consecuencias éticas.

Deberíamos preguntarnos seriamente con qué concepto de "inteligencia" nos estamos manejando o, más específicamente, qué tenemos como horizonte de sentido cuando decimos "Inteligencia artificial": ¿a qué inteligencia nos referimos? ¿qué inteligencia emula o supera la IA?

A principios del siglo XX el concepto de "inteligencia" era monolítico y factorial, es decir, pensadores como Binet entendían

que la inteligencia tenía que ver con la capacidad y la rapidez del aprendizaje. A partir de este modelo capacitista surgen los test de medición de inteligencia y las estandarizaciones binarias que discriminan entre lo "normal" y lo "anormal". Este modelo, en una de sus variantes, contempla la posibilidad de diferentes factores intervinientes en lo que se llama un "comportamiento inteligente". Posteriormente, este análisis fue superado por un conjunto de estudios que podríamos denominar "cognitivistas", que trascienden la instancia del aprendizaje como testeo del nivel de inteligencia y comienzan a interrogarse acerca de los mecanismos a través de los cuales la mente almacena, procesa y registra la información y cómo se realiza este proceso. Piaget, Vigotsky, Brunner, por citar solo algunos exponentes de esta línea, entienden a la inteligencia como un producto social en el que el aprendizaje solo crea una zona de desarrollo potencial que se activará o no en relación con la interacción social.

Sin embargo, la inteligencia es un conjunto de procesos mentales orientada a la resolución de problemas, y por eso no puede explicarse con el modelo

monolítico (capacidad de aprendizaje) ni con el correcto funcionamiento de una estructura cognitiva. Existen diversos factores que también impiden o coadyuvan al desarrollo de las personas: académico, profesional, social, emocional, psicológico, afectivo, etc. Para la resolución de problemas, entonces, no alcanzan la lógica formal, la capacidad de abstracción, el almacenamiento y puesta en juego de complejos conocimientos y datos. También hace falta creatividad, organización, motivación, actitudes, rasgos de carácter, situaciones vivenciales, oportunidades, entre otras características más ligadas al aspecto emocional del cerebro.

> Si la inteligencia es una facultad mental que permite infinidad de operaciones, ¿puede la IA emular este tipo de inteligencia?

Para finalizar este breve *excursus* sobre la inteligencia, diremos que ella es —al momento de los estudios actuales— una facultad mental por medio de la cual es posible realizar infinidad de operaciones tales como el razonamiento, la planificación, la resolución de problemas, el aprendizaje, el relacionamiento de

conocimientos entre sí y con situaciones problemáticas a enfrentar, la interacción con el ambiente y la interacción interpersonal, etc., las que permitirían un satisfactorio desarrollo de la vida. Hoy en día, también, se sabe que en estas operaciones intervienen factores cognitivos, pero también emocionales y ambos se dan en un contexto y en un ambiente.

> Si la IA constituye dispositivos capaces de "pensar", sería por demás fructífero intentar pensar, y valga la redundancia, qué estamos diciendo cuando decimos "pensar".

La pregunta que se impone, entonces, es: ¿puede la IA emular este tipo de inteligencia? Definitivamente sí sus aspectos estrictamente cognitivos. Pero si es verdad que los aspectos emocionales colaboran con los cognitivos para lograr el desarrollo satisfactorio de la vida y ambos dependen del ambiente, deberemos decir, entonces, que la IA, esa "máquina sin alma" que tanto ha trabajado la literatura de todos los tiempos, todavía se encuentra en estadios primarios de su desarrollo, a pesar de su crecimiento exponencial.

Podríamos complejizar incluso más la cuestión si nos interrogáramos acerca de qué le aporta el aspecto "emocional" a la inteligencia. Dijimos más arriba que le aporta humanidad y, por tanto, capacidad de yerro. Es una de las distópicas posibilidades y es una de las formas de leer este escenario. También, del otro lado, podríamos pensar que el aspecto emocional, y por aportarle humanidad, le agrega límites éticos que la máquina, por más hiperinteligente que sea puede no contemplar en su repertorio de herramientas. Lo planteamos como discusión abierta. No como respuesta concluyente.

Concomitantemente, se dijo más arriba que la IA constituye dispositivos capaces de "pensar". Tal como hicimos antes, sería por demás fructífero intentar pensar, y valga la redundancia, qué estamos diciendo cuando decimos "pensar", porque ni la filosofía ni la psicología han logrado encerrar el concepto en tres o cuatro parámetros claros y distintos, como diría René Descartes. Pensar es una actividad de la mente, y ocurre en una instancia interna, fruto de un complejo entramado que se sostiene desde aspectos físicos,

emocionales, cognitivos y espirituales que incluye a la persona toda.

Más allá de estas generalidades algo vagas, hay muchas especificidades diferentes según quién las formule y desde qué ámbito. En este sentido, decir que una máquina "piensa", debemos saberlo de antemano, estará también sujeto a esta abstracción y generalización dentro de la cual cabe todo o puede no caber nada.

Volvamos a la Inteligencia Artificial y resumamos lo antedicho. La IA busca desarrollarse, básicamente sobre cuatro líneas fundacionales que intentaremos resumir, siguiendo a Nils John Nilsson[22]:

> *1.* La primera tiene que ver con lo que Nilsson llama "búsqueda heurística". La resolución de problemas, hasta el advenimiento de la IA, se realizaba con operaciones numéricas y alfanuméricas a través de algoritmos cuyos programas debían contener el conocimiento completo del caso a resolver. Sin embargo, no se

[22] (1933-2019). Científico estadounidense especialista en Ciencias de la computación por la Universidad de Stanford.

tardó mucho en advertir que la mayoría de los problemas más interesantes que se presentan a diario no podían reducirse a números. Los seres humanos, mientras tanto, demostraron poseer capacidad de advertir vías de solución que las máquinas no conciben en su estructuración numérica. Reducir el área de búsqueda, por ejemplo, es un proceso cognoscitivo complejo que los seres humanos realizamos con cierta facilidad, partiendo de un abanico inmenso de posibilidades que iremos descartando naturalmente para tomar decisiones. Para imitar este comportamiento humano, Nilsson propone la heurística, que es una estrategia de simplificación que funcionaría de la siguiente forma (Hidalgo s/f): de entre el conjunto de "estados" posibles, encontrar el estado requerido para la solución del problema. A este proceso se lo denomina "búsqueda en el espacio de estados", siendo el "espacio de estados" el conjunto de estados que contiene un problema y las posibles acciones del agente. Difiere de otros métodos de búsqueda en que los nodos se generan

en el momento de la exploración y son descartados después. Los "estados" incluyen todas las configuraciones posibles, desde los "estados iniciales" hasta los "estados finales" (la solución del problema). El conjunto de "estados" forma un "grafo", que es un conjunto de datos utilizado para el modelado de un problema. La resolución del problema se obtiene mediante la proyección de las diferentes acciones del agente. La solución computacional a un problema consiste en encontrar un camino desde el *estado inicial* del *espacio de estados*, hacia un *estado objetivo.* Dentro de este paradigma, se conoce como "operadores" a un conjunto delimitado de acciones que transforman unos estados en otros.

2. Por otro lado, la creación de redes neuronales artificiales (ANN, Artificial Neural Networks) que imitan a las del cerebro, aunque todavía estén lejos de lograrlo. Muy tempranamente, a comienzos del siglo XX, Santiago Ramón y Cajal descubrió que las neuronas eran células cerebrales individuales e independientes, que se

conectaban entre sí y de esta forma enviaban información entre ellas. La conexión se establecía a través de los "axones" y las "dendritas" y, aunque no podía verlo Ramón y Cajal, planteó la hipótesis de la "sinapsis" como ese "hueco" entre neuronas donde presumiblemente se aloja el pensamiento, el conocimiento, los recuerdos, y desde allí, la información se dispersa por toda la "red neuronal". Con esta teoría ("Teoría neuronal") se oponía a la de su compañero de Nobel (1906) que entendía el funcionamiento del cerebro como una compleja red reticular (Ramón y Cajal, 1954). Dijimos más arriba que, todavía, la IA está lejos de lograr emular exactamente el funcionamiento del cerebro. Basta un ejemplo para graficar por qué estamos diciendo esto: el cerebro cuenta con miles de millones de células nerviosas, organizadas en patrones que se ocupan de algo específico del funcionamiento general del cuerpo humano, incluidas las "tareas" que no son estrictamente "físicas" como pensar o amar. No obstante, las redes neuronales artificiales serían capaces de

aprender de la experiencia, generalizar tomando como base ejemplos previos, y realizar abstracciones en base a un conjunto de datos. La diferencia entre una computación tradicional y una computación neuronal reside en que la primera programa los algoritmos mientras que la segunda "entrena" los algoritmos, y por lo tanto estos serán capaces (a imitación del funcionamiento neuronal del cerebro humano) de realizar innumerables operaciones.

3. Otra línea de la que se esperan buenos resultados es la de los llamados "algoritmos genéticos", pensados en base al código genético humano, y a las cuestiones de la evolución y la herencia. Es una técnica que tiene como presupuesto el fenómeno de la selección que prevalece en la naturaleza, de acuerdo con el cual, supuestamente, los individuos más aptos son los que sobreviven, al tener la facultad de adaptarse mejor al entorno, aun en contextos hostiles. Al día de hoy se sabe que estos cambios se registran a nivel genético, y es por eso, también, que se

transmiten como herencia. En la década del 60, John Henry Holland (1975) tuvo la idea de estudiar los mecanismos de evolución y adaptación de la naturaleza y extrapolar los resultados a los sistemas computacionales. Así sentó las bases de la "Computación evolutiva", aplicando esta emulación de la naturaleza a la resolución de problemas. La idea que subyace al planteo de Holland es que para cada problema no debe buscarse una solución óptima, sino una "población de soluciones" que, mediante el cruzamiento y operadores de recombinación, emulará la "selección natural" y llegará a la mejor solución de entre muchas. El algoritmo genético se desplaza desde una población de "cromosomas" (bits) a otra, a través de cruces, mutaciones e inversión, como en la genética. Los algoritmos genéticos, entonces, son algoritmos de búsqueda que trabajan con la supervivencia de los más compatibles, en una estructura de información. Lo novedoso es que los AG trabajan sobre la información histórica, pero con miras a mejorar el

funcionamiento de los puntos de búsqueda (Goldberg, 1989). La aplicación de los AG es prometedora: desde la resolución de problemas matemáticos abstractos, clasificación de patrones, optimización de líneas de montaje, predicción en la bolsa de valores, ingeniería aeroespacial, diseño de microchips, biología molecular, diseño de horarios de aeropuertos, entre muchas otras aplicaciones.

4. Por último, el logro de razonamiento según las reglas de la lógica formal. La lógica es la disciplina que tiene como objeto el análisis de los razonamientos para verificar si son válidos o inválidos, correctos o incorrectos.

Los razonamientos están compuestos de proposiciones, tales que algunas son premisas y otras conclusiones. Existen reglas lógicas para verificar si el razonamiento es válido: principio de identidad, principio de no contradicción, *modus ponens*, *modus toll*ens. La lógica formal no garantiza la verdad o falsedad del razonamiento,

sino si está construido correctamente según las leyes de la lógica, es decir si es lógicamente válido.

Podría decirse que la Inteligencia Artificial hace su aparición en una conferencia en el departamento de matemática de la Universidad de Darmouth en 1956 cuando John McCarthy pronunció la célebre frase: *"Este estudio procederá sobre la base de la conjetura de que todos los aspectos del aprendizaje o cualquier otro rasgo de la inteligencia pueden, en principio, ser descriptos de una forma tan precisa que se puede crear una máquina".*

> Todos los actores sociales y de los ámbitos académicos deberíamos preguntarnos cómo será la vida y cuál será su sentido a partir de las transformaciones tecnológicas que se avecinan, pero, y, sobre todo, cómo se redefinirán las reglas de juego interpersonales, las interacciones, las decisiones y las comunicaciones en estos nuevos contextos.

Desde entonces, esta predicción está generando un gran entusiasmo y a su vez, también, ciertas preocupaciones: esta tecnología emergente se acerca a gran velocidad y proyecta un mundo de posibilidades

futuras en las acciones y la convivencia de los seres humanos en esta nueva disrupción tecno-socio-digital. Bajo cierto liderazgo de inteligencias sintéticas (IS) se imponen gradualmente y de forma muy sigilosa —como mencionamos anteriormente— en todas las áreas de nuestras sociedades.

Claro está que estos escenarios y tecnologías de inteligencia artificial no deberían emerger sin un profundo análisis, cuestionamiento y reflexión crítica, desde la filosofía, la sociología, la pedagogía, la psicología e incluso la antropología. Todos los actores sociales y de los ámbitos académicos deberíamos preguntarnos cómo será la vida y cuál será su sentido a partir de las transformaciones tecnológicas que se avecinan, pero, y, sobre todo, cómo se redefinirán las reglas de juego interpersonales, las interacciones, las decisiones y las comunicaciones en estos nuevos contextos en los que también estarán jugando las mega corporaciones que manejan las implicancias interactivas sociales, y lo hacen de acuerdo a intereses y sesgos no siempre nobles y humanitarios.

Aparecen nuevos senderos que se bifurcan provocando una marcha social muy diferente a la que veníamos transitando, puesto que la inteligencia artificial pone en un primer plano de discusión la permanencia del concepto de "inteligencia" tal como lo conocemos hasta ahora. Estos nuevos territorios de la IA pueden ser identificados ya, a través de sus diversas propiedades y características, como agentes sintéticos y artificiales que comienzan a entrelazarse con los principales aspectos de la vida de las personas.

Dicho todo esto, la pregunta que no podemos dejar de hacernos es ¿cómo puede afectar el uso y combinación de IA a la evolución de los seres humanos? Sin lugar a dudas, el interés en responderla radica en que en no mucho tiempo nos encontraremos con sociedades que estarán muy involucradas en el desarrollo, configuración e implementación de estos agentes (IA e IS). Por lo tanto, estas generaciones y las que vienen —sin hipotetizar demasiado— tendrán como moneda corriente prácticas, formas y estilos de vida vinculados a contextos de interacción bio-artificiales inteligentes, y dinámicas de socialización, de

estudios, laborales, comunicacionales, entre otras, muchas veces gestionadas por estas nuevas formas de inteligencias. Por eso decimos que es necesario reflexionar sobre el punto para que dichas gestiones que podamos delegar en ellas no se conviertan en una nueva e invisible "inquisición algorítmica".

Por lo tanto, la prioridad es ocuparnos con urgencia inteligente sobre este tema y su prospectiva de la vinculación humana y tecnológica con este tipo de herramientas. El aspecto que debemos atender prioritariamente es el de los instrumentos que podamos brindar a las personas —y también a los responsables educativos— en cuanto a información, capacitaciones, investigaciones y prácticas críticas y reflexivas sobre su uso y apropiación comprensiva. Y todo ello, desde una perspectiva que nos permita pensar de forma ética y transparente sobre el futuro de la sociedad humana. Es que solo así podremos intentar acercarnos a sociedades que no solo deberán repensar todo el accionar de las personas, sino que también deberán atender a conceptos neurálgicos como la descentralización de la información y el conocimiento, la vinculación en red con sopor-

tes de inteligencia artificial y sintética, basados en analíticas de aprendizajes. Asimismo, también deberán entrar en el análisis los modelos y estilos de vida pensados y diseñados en cómo aprende el cerebro, la utilización de dispositivos de interacción, comunicación, enseñanza, aprendizaje y evaluación —sostenidos y potenciados por un despliegue de avatares y hologramas digitales especializados en aprendizaje profundo y experto en temas humanos—. Como los campos de conocimientos son amplios y múltiples, la reflexión deberá hacerse extensiva a las grandes ventajas y desventajas que la IA aportará a las nuevas formas de gestión y control sobre la vida y las acciones de los seres humanos.

¿Qué tan inteligente es la inteligencia artificial?

Llegados a este punto, vale puntualizar nuevamente algunos conceptos referidos a la IA, ya presente en nuestras vidas, la cual posee y proyecta, por el momento, algunos aspectos de inteligencia como la que tenemos los seres humanos.

Sobre la inteligencia Artificial (IA) actualmente se tiene la creencia de una atribución —casi mágica— de ser tecnologías superinteligentes y que los humanos ni siquiera cuentan con la posibilidad de competir frente a ellas.

Pero si investigamos un poco más en profundidad, nos daremos cuenta de que la IA —o bien los algoritmos generativos, y en especial la que hoy hemos comenzado a conocer y operar— no tienen la inteligencia de los humanos, debido a varias razones fundamentales:

1. Diferencias en la estructura y funcionamiento

La base de la inteligencia humana se halla en el órgano más complejo de las personas, es decir, en nuestro cerebro. El mismo es una gran y extensa red de neuronas, conectadas por miles de millones de sinapsis y sistemas con una complejidad que excede, hoy en día, largamente a lo que conocemos sobre él.

Por otro lado, los avances de la IA, intentan de manera muy rápida lograr avanzar sobre el desarrollo de neuronas artificiales, inspiradas en el modelo real de las neuronas del cerebro humano, siendo esta réplica

uno de los pilares más fuertes en el avance de la IA de los próximos años. Si bien las neuronas artificiales intentan emular a las neuronas cerebrales, con un conjunto de "entradas" y "salidas" hacia otras neuronas que se multiplican, sin embargo, es cierto, también, que las artificiales no tienen, por ejemplo, los retardos y sincronismos en la generación de la salida. Técnicamente, la neurona artificial es una unidad procesadora que posee cuatro elementos: el receptor, al que llegan las señales de entrada, el elemento sumador, que pondera esas señales de entrada mediante una operación algebraica, el elemento de función activadora, que dispara una salida o no, y el elemento de salida, que produce la señal y la salida de la neurona (Lara Rosano, s/f).

> El aspecto que debemos atender prioritariamente es el de los instrumentos que podamos brindar a las personas —y también a los responsables educativos— en cuanto a información, capacitaciones, investigaciones y prácticas críticas y reflexivas sobre su uso y apropiación comprensiva.

Con todo, nuestras redes neuronales biológicas poseen un altísimo nivel de complejidad y sofisticación en su estructura y formas de *conexionamiento* (conexión y relacionamiento) con el resto de los sistemas de nuestro cuerpo humano, que por el momento guarda más secretos que respuestas.

2. Aprendizaje basado en experiencias

Las personas vamos adquiriendo, desde nuestros primeros años de vida y luego en el desarrollo posterior, diversos tipos de conocimientos, habilidades y competencias, muchas de ellas basadas en la vivencia de experiencias e interacciones con el entorno, relacionadas con el sistema de activación reticular ascendente[23]. Nuestro cerebro aprende por medio de acciones, como conocer-comprender-decidir o sentir-pensar-actuar.

[23] Es un conjunto de neuronas y sinapsis interconectadas a lo largo del tronco encefálico que interviene en la regulación del sueño y de la vigilia.

Por su parte, la inteligencia artificial se nutre en su desarrollo primario a partir de sus bases algorítmicas y sobre modelos matemáticos predefinidos, los cuales se vinculan con un mundo de datos que son procesados de forma computacional, y este funcionamiento es muy distinto y limitado si lo comparamos con la forma humana. En el procesamiento del cerebro humano intervienen, también, factores emocionales, por lo tanto, la IA parte con una desventaja, al no tener esta capacidad experimental asociada a formas de comprensión del mundo y sus contextos.

3. Comprensión por contextos

La principal función del cerebro humano es la de relacionarnos con el entorno y permitirnos interactuar: conocerlo y comprenderlo para poder sobrevivir en él. El desarrollo de la inteligencia biológica humana y sus factores comprensivos en relación al contexto, no posee las limitaciones de un procesamiento de información y/o tareas acotadas y estructuradas, que le dificulten conocer y comprender la complejidad del mundo real como es el caso de la IA. Por el contrario, las formas de

relacionarse con la información se producen de manera más holística para los humanos, es decir, ante determinados eventos, el cerebro busca identificar patrones asociados al propio evento y su contexto con base a recuerdos registrados en memorias, y los interpreta en función de emociones asociadas, posibles amenazas, sentidos, para dar luego lugar a formas de comprensión y reacción basadas en funciones ejecutivas del cerebro.

> Entre la IA y la inteligencia humana hay diferencias en la estructura y funcionamiento.
> La humana desarrolla un aprendizaje basado en experiencias, realiza comprensión por contextos, tiene consciencia y emociones, y, además, creatividad y pensamiento abstracto.

4. Conciencia y emociones

Los humanos tenemos conciencia de nosotros mismos y de nuestro entorno. También experimentamos una amplia gama de emociones y sentimientos que influyen en la toma de decisiones y comportamientos.

Por el momento, la IA carece de conciencia y emociones, ya que se basa en algoritmos, y en su procesamiento de datos y analíticas[24] de tipo descriptivas, diagnósticas, predictivas y prescriptivas, sin ningún tipo de comprensión subjetiva semejante a la de la experiencia humana.

5. Creatividad y pensamiento abstracto

Los seres humanos tenemos la capacidad de pensar de manera abstracta, hacer conexiones creativas, resolver problemas complejos y generar nuevas ideas e innovaciones. En este sentido, la IA viene logrando avances en la generación de contenido y en tareas creativas, aunque por el momento todavía está limitada en términos de originalidad y pensamiento en comparación con los seres humanos.

---

[24] La IA es capaz de analizar grandes volúmenes de datos de forma masiva y automática y, por esta razón, forma parte de lo que se conoce como "Data Mining" o "Minería de datos" y "Big data" (grandes volúmenes de datos) y, a su vez, y de forma más amplia, de lo que se conoce como "Ciencia de datos". Mediante esa automatización de procesos algorítmicos, la IA es capaz de ofrecer soluciones a problemas que se le planteen.

Es decir, aunque la inteligencia artificial ha avanzado significativamente en muchas áreas y puede superar a los humanos en tareas específicas, tanto en sus dimensiones de velocidad como en su capacidad de procesamiento, todavía está lejos de replicar la complejidad y diversidad de la inteligencia humana en su conjunto. La inteligencia artificial actual carece de la experiencia, de conciencia, de emociones, de creatividad y de la comprensión profunda que caracterizan a la inteligencia humana.

En resumen, podemos definir a la inteligencia artificial (IA) como un campo de estudio, de investigación y de promoción de desarrollos abocados a crear sistemas capaces de emular la inteligencia humana en sus procesos mentales y de razonamiento, pero también en sus aspectos conductuales. Pensar como humanos sería una de las metas de la IA, sin embargo, choca con la dificultad de no poder, al día de hoy, especificar sin margen de duda cómo pensamos efectivamente los humanos. En todo caso, lo que intenta la IA es simular los procesos de cognición humana que

hasta ahora se conocen o han sido estudiados. Paralelamente, otro intento es que la IA "piense" con las reglas de la lógica formal, tal como explicamos antes, a fin de que produzca razonamientos válidos. Todavía estamos en este sentido en el terreno de las aspiraciones. Con todo esto, lo que se espera es que la IA se convierta en un "agente racional", es decir, que no solo piense de acuerdo con la lógica, sino que actúe obteniendo las inferencias correctas, para llegar a las conclusiones lógicas a fin de resolver los problemas.

Dicho todo lo cual hemos de considerar que, aunque la IA ha logrado avances significativos en diversas áreas, todavía hay una brecha muy importante entre la inteligencia digital de las máquinas y la de los seres humanos. Hay varias razones clave por las que la IA no posee la inteligencia biológica de los humanos:

1. Naturaleza de la inteligencia humana

La inteligencia humana es el resultado de una interacción compleja entre el cerebro, la mente, los sentidos, las emociones y la experiencia acumulada a lo largo de la vida. Aunque se han logrado avances en el

desarrollo de redes neuronales artificiales y otros enfoques de IA, estos sistemas no son capaces de replicar completamente la forma en que los humanos procesamos y comprendemos la información.

2. Conciencia y subjetividad

La inteligencia humana no solo consiste en procesar información y tomar decisiones lógicas, sino también en ser conscientes de nuestro entorno, nuestras experiencias y nuestras propias emociones. Hasta ahora, la IA carece de conciencia y subjetividad. Los sistemas de IA no poseen emociones y no tienen una experiencia personal que los moldee. Esto limita su capacidad para comprender completamente el mundo y responder de manera significativa.

3. Flexibilidad y adaptabilidad

Los humanos somos inherentemente flexibles y adaptables en la forma de pensar y actuar, gracias a la neuroplasticidad de nuestro cerebro. Podemos aprender nuevas tareas y adaptarnos a diferentes situaciones de manera rápida y eficiente. En cambio, los

sistemas de IA están diseñados para tareas específicas y requieren de entrenamiento y programación para desempeñarse en un ámbito determinado. Aunque pueden ser muy buenos en tareas específicas, su capacidad para generalizar y adaptarse a situaciones nuevas es limitada en comparación con la de los humanos.

Por todo lo que venimos exponiendo, podemos advertir que las posibilidades que se abren de cara al futuro de la mano de la IA son múltiples, diversas y prometedoras en cuanto a mejorar casi todas las áreas de la vida humana en este planeta. Las aplicaciones de la IA son incontables: en la medicina, la ingeniería, la robótica, en interfaces y sistemas de ayuda, en educación, en gestión de información, en matemáticas, en la industria del

> Humanizar la tecnología parece un sinsentido, sobre todo para aquellos que pretenden tecnificar la humanidad, mejorarla, trascenderla por medio de la IA, de los metaversos, y de toda la gama de tecnoherramientas que tenemos al presente.

entretenimiento, en la biología, en lo jurídico, en la arquitectura, el diseño urbano, en literatura, arte y música, en el comercio, en la ingeniería aeroespacial, en las actividades militares, en fin, tan lejos como pueda llegar nuestra mirada podrá extenderse el área de acción de la IA.

Contar con una herramienta que reduzca al máximo los errores humanos en cualquier área es una aspiración lícita con la que han soñado no solo los científicos sino también los utópicos de cualquier disciplina humana. Ahora bien, la pregunta que podríamos hacernos, y con la cual queremos terminar este capítulo es: ¿la IA es igual (o aspira a ser) o semejante a lo humano menos el error humano? ¿Esa sería la ecuación? Algo así como:

IA = ~humano-error humano

O quizás podríamos pensar que aquello que resulta en *error* muchas veces es producto de un conglomerado holístico de características biopsicosociales y espirituales inherentes al ser humano —y al ser hu-

mano en sociedad— como las emociones, la consciencia ética, la responsabilidad social o moral, las elecciones personales, la ideología, nuestras creencias e incluso los gustos, las tendencias, las inclinaciones, los rasgos de personalidad, en fin, aquello que nos hace, justamente, humanos y no seres maquinales. Queremos decir: ¿De qué estamos hablando cuando hablamos de error? ¿Hay algo de lo objetivo en la concepción del "error"? ¿O puede ser que lo que para alguien sea error para otro u otra sea un acierto?

No pretendemos en estas cortas líneas ahondar en toda la complejidad que un tema como este plantea. Simplemente, sí queremos llamar la atención a dicha complejidad. Y esa complejidad nos motiva a proponer ya, cuando todavía la IA no ha obtenido todas sus potencialidades, una Inteligencia Artificial Humanizada.

Sabemos que habrá quien vea una tautología en este sintagma, toda vez que la IA es una emulación de la inteligencia humana (o pretende serlo). Pero no es realmente una tautología, porque lo que proponemos es que el elemento humano, en vez de ser una limitante sea un plus de valor, un valor agregado que proporcione un

marco regulatorio ético. Esto es que, así como existe —por fuerza de las circunstancias— una bioética que intenta encarrilar las prácticas científicas con seres humanos, así también haya una ética aplicada a la IA que incluso regule legalmente sus usos, alcances y aplicaciones.

Humanizar la tecnología parece un sinsentido, y mucho más, intentar lograr una tecnología del alma, sobre todo para aquellos que pretenden tecnificar la humanidad, mejorarla, trascenderla por medio de la IA, de los metaversos, y de toda la gama de tecno herramientas que tenemos al presente.

Sin embargo, y tal vez, humanizar este mundo —incluso a través de las máquinas— sea lo que nos preserve como género humano y conserve para nosotros y nosotras y las nuevas generaciones la casa común en la que todos habitamos.

Referencias bibliográficas

Acosta Flores, J. (2016). Ingeniería de sistemas. Un enfoque interdisciplinario. Alfaomega.

Alvarado Rojas, M. (2015). Una mirada a la inteligencia artificial. *Rev. Ingeniería, Matemáticas y Ciencias de la Información.* Vol. 2 / Núm. 3 / enero - junio de 2015; pág. 27-31.

Angulo Rojas, R. (2021). *Algoritmos genéticos. Laberintos & Infinitos*. Recuperado de: http://laberintos.itam.mx/algoritmos-geneticos/

Barrera Arrestegui, L. (2012). Fundamentos históricos y filosóficos de la inteligencia artificial *UCV-HACER. Revista de Investigación y Cultura*, vol. 1, núm. 1, julio-diciembre, 2012, pp. 87-92

Castro, V., & Alejandra, C. (2015). Revisión histórica del concepto de inteligencia. Recuperado de: https://www.semanticscholar.org/paper/Revisi%C3%B3n-hist%C3%B3rica-del-concepto-de-inteligencia-Castro-Alejandra/ee3076755163467e0a3fc6b120c83e710057b70f

Di Tada, E. (2002). Algoritmos genéticos. *Revista ciencia y tecnología*. Universidad de Palermo. Recuperado de: http://dspace.palermo.edu/dspace/bitstream/handle/10226/1395/Art%202%20CyT%202%20Algoritmos%20Gen%c3%a9ticos%20di%20Tada.pdf?sequence=1&isAllowed=y

Goldberg, D. (1989). *Genetics Algorithms in Search, Optimization and Machine Learning*. Addison Wesley.

Haugelman, J, (1985). *La inteligencia artificial*. Siglo XXI.

Hidalgo, L. (s/f). *Inteligencia artificial y sistemas expertos*. Universidad de Córdoba. Recuperado de file:///D:/Users/W10/Downloads/Luis%20Amador_Inteligencia%20artificial_1996-1.pdf

Holland, J. (1975). Adaptation in natural and artificial systems. An introductory análisis with applications to biology, control, and artificial intelligence. The university of Michigan Press.

Lara Rosano, F. (s/f). *Fundamentos de redes neuronales artificiales.* Laboratorio de cibernética aplicada. UNAM. Recuperado de: https://conceptos.sociales.unam.mx/conceptos_final/598trabajo.pdf

McCarthy, J. (1955). A proposal for the Dartmouth Summer Research Project on artificial intelligence. Recuperado de:http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

McCarthy, J. (2004). What is AI? / Basic Questions. Recuperado de: http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html

Nilsson, N. (1987) Principios de Inteligencia Artificial. Ediciones Díaz de Santos.

Nilsson, N. (2001). Inteligencia artificial. Una nueva síntesis. Mc Graw Hill

Ramón y Cajal, S. (1954). Neuron Theory or reticular theory? Objective evidence of the anatomical unity nerve cells. Consejo Superior de Investigaciones Científicas.

Rodriguez de Romo, A. (2005). ¿Reticularismo o neuronismo?; diferente percepción de la misma circunstancia. Arch Neurocien (Mex). Vol. 10 No. 1: 26-32, 2005

Rodriguez Piñero, P. (s/f). Introducción a los algoritmos genéticos y sus aplicaciones. Recuperado de: https://www3.gobiernodecanarias.org/medusa/ecoblog/rgonfal/files/2012/04/J24C.pdf

Russell, S. Norvig, P. (2004). Inteligencia artificial. un enfoque moderno. Pearson Educación.

Salmerón Vílchez, P. (2002). Evolución de los conceptos sobre inteligencia. Planteamientos actuales de la inteligencia emocional para la orientación educativa. Educación XXI, núm. 5, 2002, pp. 97- 121

Segovia, A. (2000). El pensamiento: una definición interconductual. Revista de Investigación en Psicología, Vol.3 No.1, Julio 2000

Tolmos, Piedad. (2002). Introducción a los algoritmos genéticos y sus aplicaciones. Rect@. Actas_10. Recuperado de: https://www.researchgate.net/publication/26440754_Introduccion_a_los_algoritmos_geneticos_y_sus_aplicaciones

Verrastro, C., Gómez, J., Alcoberro, R. (2009). Algoritmos genéticos. GIAR. Recuperado de: https://www.frba.utn.edu.ar/wp-content/uploads/2021/02/AlgoritmosGeneticos-compressed.pdf

Capítulo 5

Prometeo Digital
El nuevo fuego de los humanos

PROMETEO. Sí. Hice que los mortales dejaran de pensar en la muerte antes de tiempo.
CORIFEO. ¿Qué solución hallaste a este mal?
PROMETEO. Albergué en ellos esperanzas ciegas.
CORIFEO. Gran favor otorgaste a los mortales.
PROMETEO. Además de esto, yo les regalé el fuego.
CORIFEO. ¿Y ahora los efímeros tienen el fuego resplandeciente?
PROMETEO. Por él aprenderán muchas artes.

Cuando los homínidos comenzaron a dominar el fuego, hace aproximadamente 1.5 millones de años, ese mismo fuego que ya conocían pero ignoraban cómo producirlo y cómo ponerlo a su favor, efectivamente hubo un cambio de paradigma que hizo dar un salto cualitativo a la humanidad: dominar el medio ambiente, mejorar las condiciones de vida, cocinar la carne —hecho no menor que trajo como consecuencia incluso cambios biológicos en el aparato masticatorio y digestivo, aunque también una mejora en la alimentación con sus consecuencias benéficas para la salud—, cambios en los conglomerados humanos, en

sus condiciones habitacionales, transformaciones en los métodos de caza y de trabajo a través de la fabricación de herramientas y armas, seguridad, etc.

Como toda transformación inesperada, y en épocas en que el pensamiento racional y el estado del conocimiento no estaban desarrollados como actualmente, cada civilización explicó este acontecimiento con ribetes míticos y milagrosos diferentes.

> El bien, para cualquier ser humano, tiene una contracara de mal: lo que es bueno, en manos humanas, puede también ser malo.

En el caso de la tradición griega, el mito de Prometeo es uno de los elegidos para dar cuenta de la historia pasada: un Titán que roba el fuego a Zeus y lo ofrenda a los seres humanos y, en su comportamiento infatuado, es castigado a estar encadenado y hostigado por aves que le comían el estómago y el corazón, acaso para un griego, el asentamiento de la persona, de lo que hoy llamaríamos "alma", de la vida misma, del ser.

¿Por qué Zeus castigó a Prometeo, si lo que había hecho era en apariencia bueno, benéfico para la humanidad toda? La respuesta no es sencilla. Con toda seguridad no es por "robarlo" o por "desobedecer" a Zeus… Tal vez —pero solo tal vez— Zeus considerara que los seres humanos no estaban preparados para ese salto cualitativo que el fuego les hizo dar. Otro mito, esta vez semita, nos ayuda a comprender la idea: había algunos árboles en ese maravilloso huerto de Edén, pero había uno "de la ciencia del bien y del mal" que estaba en el centro (justo en el centro, lugar bien simbólico). Comer de este árbol implicaba la muerte. Era incluso más grave que comer del árbol de la vida y tener la pretensión de vivir para siempre. ¿Por qué? Porque entrañaba un cierto peligro —y peligro de muerte—. Es que el bien, para cualquier ser humano, tiene una contracara de mal: lo que es bueno, en manos humanas, puede también ser malo. Por eso no hay dos árboles, uno de bien y uno de mal. El bien y el mal, lo bueno y lo malo son parte consubstancial del ser humano. Por eso es tan grave el acto insurrecto de comer del fruto, cuando quizás no estaban preparados para hacerlo. No era malo

en sí mismo: ¿pero con qué intención comían el fruto? En general, nada es malo en sí mismo, pero ¿con qué intenciones se usa o con qué intenciones se hace? Comer del árbol significaba, en el mito, pretender ser como Dios, pero, consecuentemente, sin las posibilidades de abstenerse del mal: querer ser Dios aun cuando nuestro rostro es jánico, imperfecto, ambivalente (Valzura, 2023).

Es posible que en esta explicación hallemos un paralelismo con el desarrollo de las tecnologías en general y de la IA en particular: son *buenasmalas*, y están en el centro de este momento de la historia. ¿Qué hacer? ¿Dejarlas en el "árbol" para no causar daño? No parece posible, toda vez que el desarrollo ya es imparable, y no parece apropiado desdeñarlo, por la gran cantidad de beneficios que nos aporta. Quizás sería mejor estar prevenidos: todo lo bueno puede ser malo, según cómo se utilice, para qué se utilice y quién lo utilice. La vida y las elecciones, parece decir el mito genésico, tienen un riesgo que asumir. Vivir es un riesgo. Aunque los riesgos no se afrontan con displicencia o resignación, sino con ojos bien abiertos y previsiones, y

de eso estamos hablando. Como en el mito genésico, Prometeo sufrió las consecuencias de pretender manipular el fuego sagrado sin pensar en las consecuencias. Así logró explicar la humanidad de entonces por qué algo —el fuego— que parecía bueno sin más, importante para la evolución, crucial para un cambio absoluto de paradigma histórico, también había traído aparejada una cantidad de problemas no previstos. Y eso es el mito, finalmente, un intento de explicar la narración de la historia.

Naturalmente los seres humanos desarrollamos prácticas cognitivas "en espiral", una suerte de "preguntas" que se disparan en nuestra mente y cerebro con las que siempre buscamos detectar patrones que nos permitan identificar amenazas. Por eso mismo, frente al desafío de la IA, nos encontramos de forma constante analizando e intentando reflexionar cómo los avances —en especial de las tecnologías emergentes— pueden ayudarnos en nuestra cotidianeidad, sin que esto represente un problema o un perjuicio para nuestra vida, tanto individualmente como en la práctica social.

Indudablemente, la inteligencia artificial, y sus posibles derivaciones buenas y malas, son objeto de curiosidad y preocupación no solo de los filósofos, sociólogos y antropólogos, sino también de la gente común que desea estar informada. Creemos que, en especial, los y las docentes miran esta tecnología con entusiasmo y desconfianza a la vez. Así lo atestiguan las tendencias en los motores de búsquedas, que dan cuenta de que esas búsquedas se cuadruplicaron en los países de la región, y aumentaron casi un trescientos por ciento en Europa. La corporación internacional Capgemini, dedicada a la consultoría en tecnología, expone que la mitad de los docentes consultados en un estudio de campo sostienen la necesidad de cambiar los planes de estudio para mejorar su actividad docente mediante la incorporación de la IA, sin embargo, un 78% se muestra preocupado por un posible impacto negativo en los resultados del aprendizaje, en la pérdida de las habilidades de escritura y en la limitación de la creatividad.

Tal como en el mito de Prometeo, la IA nos plantea interrogantes y cuestionamientos acerca de su

uso y desarrollo: una nueva fuerza que desafía los límites establecidos, como el fuego de los dioses, ese que no tenía un destino para los humanos, pero ya en sus manos les otorga conocimientos muy significativos para su desarrollo, y acaso también peligrosos. Esquilo le atribuye a Prometeo haberles enseñado a los seres humanos todo lo necesario para evolucionar:

> Al principio, miraban sin ver y escuchaban sin oír, y semejantes a las formas de los sueños en su larga vida todo lo mezclaban al azar. No conocían las casas de ladrillos secados al sol, ni el trabajo de la madera; soterrados vivían como ágiles hormigas en el fondo de antros sin sol. No tenían signo alguno seguro ni del invierno, ni de la floreciente primavera ni del estío fructuoso, sino que todo lo hacían sin razón, hasta que yo les enseñé los ortos y ocasos de los astros, difíciles de conocer. Después descubrí también para ellos la ciencia del número, la más excelsa de todas, y las uniones de las letras, memoria de todo, laboriosa madre de las Musas. (…) Sólo yo inventé el vehículo de los marinos, que surca el mar con sus alas de lino. Y, mísero de mí, yo que he encontrado estos artificios para los mortales, no tengo artimaña que pueda librarme de la actual desgracia. (…) Las artes y recursos que ideé. Lo más importante: si uno caía enfermo, no había ninguna defensa, ni alimento, ni unción, ni

pócima, sino que faltos de medicinas morían, hasta que les enseñé las mezclas de remedios clementes con los que ahuyentan todas las enfermedades. (…) En una palabra, sabe todo a la vez: todas las artes para los mortales proceden de Prometeo.

Hesíodo, por su parte, explica que, como castigo a su osadía, Prometeo fue atado con "irrompibles ligaduras" y condenado a que un águila grande le comiera el hígado inmortal, el que volvía a crecer durante la noche para repetir la historia de destrucción vez tras vez.

> Prometeo sufrió las consecuencias de pretender manipular el fuego sagrado sin pensar en las consecuencias. Por eso mismo, frente al desafío de la IA, intentamos reflexionar cómo los avances pueden ayudarnos en nuestra cotidianeidad, sin que esto represente un problema o un perjuicio para nuestra vida, tanto individualmente como en la práctica social.

Es que el desarrollo de una tecnología sin los conocimientos adecuados y sin una conciencia moral sobre sus usos justos, equitativos y éticos, nos puede conducir a grandes problemas y, por qué no, dolores que no supimos ver y mucho menos anticipar.

La Inteligencia Computacional, este fuego prometeico que el progreso nos regala como tecnología emergente, tiene un potencial formidable que permite a las personas realizar muchas tareas y funciones que antes eran desarrolladas exclusivamente por el cerebro y la mente de los seres humanos. Una de ellas es la capacidad de procesar inmensos volúmenes de datos e información y, mediante el desarrollo de *big data* y de minería de datos, analizar patrones complejos, lo que habilita la toma de decisiones basadas en arquitecturas digitales.

Las arquitecturas digitales, mientras tanto, están diseñadas con identificaciones hiperindividualizadas y geolocalizadas con datos que se proporcionan —de manera inconsciente— por vía de nuestras prótesis digitales, como los teléfonos inteligentes, dentro de los laberínticos territorios de las redes sociales.

Estos diseños, claro está, y la simbiosis que se manifiesta a través de esta transformación, tienen como finalidad no solo desarrollar una inteligencia computacional. Sin embargo, hasta el presente solo se ha alcanzado el nivel de algoritmos generativos que

comienzan a involucrarse dentro de la gran mayoría de las actividades de los humanos. No obstante, a pesar de este aparente límite, no debemos descuidarnos ni caer en miopías de la suposición, ya que las verdaderas metas de las grandes corporaciones consisten en superar estos algoritmos y así alcanzar el diseño, construcción y usos de cerebros artificiales.

El filósofo francés Eric Sadin (2018) en el libro originalmente de 2013 titulado *La humanidad aumentada*, plantea el panorama en el que actualmente estamos viviendo, como consecuencia del aumento sostenido de las tecnologías: la realidad misma ha cambiado su configuración y está siendo mediada, de manera que los estratos informacionales se superponen con nuestra directa aprehensión de las cosas, pero, además, la humanidad misma está "aumentada". En el pensamiento de Sadin ya podríamos hablar de una condición híbrida, una verdadera "antrobología" (condición humana asistida por tecnología) o la combinación de la inteligencia humana y artificial. Su intuición es inquietante: *"Tal vez, hemos penetrado sin darnos cuenta dentro de una matriz compuesta de códigos que nos engloba por completo, al*

igual que a los 'humanos' que deambulan en el plasma electrónico de la Matrix" (p.83). Las comparaciones que hace Sadin con Prometeo, el Golem o Frankenstein son, en sus palabras, "de una monstruosidad amenazante".

Las predicciones de este autor son alarmantes, y nos parece interesante recordarlas. Como dijimos antes, no para paralizarnos o dar la espalda a lo que ya está entre nosotros y a lo que muy pronto vendrá, sino para estar alertas: convivimos con una realidad paralela y con una humanidad paralela y, entonces, ¿qué lugar ocupamos los seres humanos? ¿cómo imaginamos esta convivencia?

Todavía el plano de la emocionalidad, la sensibilidad, la empatía, la compasión —en tanto padecer con—, la posibilidad de ponerse en el lugar del otro/otra, el reconocimiento, el amor, la admiración, son cualidades típicamente humanas. Si la razón carece de sentimiento es posible que no sea humana, puesto que no existe la opción de razonar —como "razón-aséptica"— sin estar atravesados y atravesadas por la propia subjetividad, por las propias emociones y sentimientos. Entonces, la mirada crítica que estamos

proponiendo sobre las tecnologías en general y sobre la IA en particular es una que está sólidamente fundamentada en lo racional-humano, es decir, en una racionalidad empática y sensible, esa que los antiguos situaban en el corazón y no en el cerebro.

Referencias bibliográficas

Esquilo (1987). *Tragedias*. Gredos.

Hesíodo (1978). *Teogonía*. Gredos.

Sadin, E. (2018). *La humanidad aumentada. La administración digital del mundo*. Caja Negra

Valzura, E. (2023). *Teología indisciplinada. Una apuesta por la teología asistemática y liberadora.* Ediciones Diapasón.

Capítulo 6

Anatomía de un *prompt*
Elementos esenciales para una comunicación eficaz con la Inteligencia Artificial

La revolución de la inteligencia artificial no radica tanto en la construcción de una economía de la comodidad, sino en la forma en que pone en juego la autonomía de lo humano.

Eric Sadin.

Algunas consideraciones preliminares

La idea central de este capítulo es presentar una modulación entre el Modelo Neuro-Tecno-Pedagógico que ya hemos desarrollado en los libros anteriores —y mencionado en este mismo libro— y la Inteligencia Artificial, en este caso particular la IA Gemini de Google. El propósito es mostrar la importancia de aplicar el Modelo en nuestras interacciones con la IA con fines pedagógicos, de manera de integrar el uso de esta novedosa tecnología con la práctica docente, pero de acuerdo a pa-

rámetros que permitan enriquecerla, conducirla por senderos más interesantes y productivos y, en consecuencia, favorecer los procesos de enseñanza, aprendizaje y evaluación.

En virtud de estos propósitos, este capítulo utilizará la IA para la exposición de ejemplos, a fin de que sea más clara y evidente la interacción que puede darse —y que proponemos es necesario que se dé— entre el Modelo Neuro-Tecno-Pedagógico que presentamos y la utilización siempre en expansión de la IA.

¿Qué es un *prompt* y por qué debería importarme?

Un *prompt* en IA es una instrucción o texto inicial que se proporciona a una herramienta de IA generativa para guiar su generación de respuestas o resultados. Es la clave que permite a los usuarios interactuar con la IA y obtener los resultados deseados. Por eso, los *prompts* son la base de la interacción con los modelos de IA conversacional y generativa. Esto significa que los *prompts* son preguntas o frases con las que se intenta interactuar con la IA. Se podría decir que son el primer

paso de una "conversación", entendida esta como una interacción con la IA. Son *prompts* las preguntas sencillas que buscan una respuesta, pero son también *prompts* las minuciosas indicaciones que damos a la IA para producir un ensayo.

Un *prompt* bien escrito puede generar resultados sorprendentes, mientras que uno mal escrito puede frustrar al usuario y obtener resultados no deseados. Esto significa que al confeccionar el *prompt* estamos influyendo directamente en la respuesta y en la calidad de la respuesta o del producto final, y por ello se deben tener en cuenta algunas premisas para sacar de los *prompts* y de la IA los mejores resultados y los más satisfactorios.

> El propósito es mostrar la importancia de aplicar el Modelo en nuestras interacciones con la IA con fines pedagógicos, de manera de integrar el uso de esta novedosa tecnología con la práctica docente, pero de acuerdo a parámetros que permitan enriquecerla, conducirla por senderos más interesantes y productivos y, en consecuencia, favorecer los procesos de enseñanza, aprendizaje y evaluación.

Un *prompt* es, básicamente, texto. De hecho, y más específicamente,

una plantilla de texto que genera texto. Sus utilizaciones son varias, desde traducir idiomas, escribir diferentes tipos de contenido creativo, hasta responder a sus preguntas de manera informativa.

Los *prompts* no solamente son indicadores para obtener información específica (informal o relevante), sino para realizar investigaciones y escribir ensayos. Pero la escritura formal es solo una parte de las amplias posibilidades que se abren en abanico delante de los usuarios: hay una emergencia infinita de creatividad que puede dispararse a partir de la propia creatividad de quienes utilizan los *prompts,* y esas posibilidades infinitas se extienden, incluso, a la obtención de contenido creativo e innovador, como poemas, historias, códigos, guiones, etc.

A continuación, exploraremos los puntos más importantes a tener en cuenta al analizar y crear *prompts* para IA.

a) Definición del objetivo

El primer paso es definir claramente el objetivo que se quiere lograr con el *prompt.* ¿Se busca obtener

información, generar contenido creativo, o simplemente entablar una conversación? Como en toda otra actividad, definir el objetivo es fundamental para el éxito o el fracaso de lo que se busca lograr. Asimismo, como en toda interacción comunicacional, se debe tener en cuenta el público objetivo al que está dirigido. Esto permitirá ajustar la profundidad del análisis, el vocabulario elegido de acuerdo con las competencias de ese público objetivo, el tono, y las restricciones al universo del discurso que ambas partes comparten o no comparten pero que, de alguna forma, rigen la comunicación.

b) Comprensión de la arquitectura de una IA

Es importante conocer las capacidades y limitaciones del modelo de IA con el que se va a trabajar. Algunos modelos son mejores para generar contenido factual, mientras que otros son más adecuados para la creatividad.

c) Claridad y concisión

El *prompt* debe ser claro, conciso y fácil de entender. Debe evitarse la ambigüedad y el uso de

lenguaje técnico o especializado, porque el *prompt* debe ser fácil de entender para el modelo. Este ítem abarca: correcto uso de terminología, coherencia y cohesión de la frase, apego a las reglas sintácticas y de la gramática, evitar, en lo posible, las anfibologías. Hay que tener en cuenta que hay expresiones que se comprenden en una conversación oral entre usuarios del lenguaje que comparten las mismas competencias lingüísticas, pero pueden ser mal entendidos por la IA.

Resumiendo:

- Sea preciso y conciso en su solicitud.
- Evite la ambigüedad y las frases vagas.
- Use un lenguaje sencillo y directo.

d) Especificidad

No solo la claridad es necesaria, también lo es la especificidad. Por ejemplo, si mi *prompt* es "escribe un ensayo", la directiva es clara: lo que quiero de la IA es un ensayo. Pero, evidentemente, no es específico: existen millones de posibilidades de ser interpretado este *prompt.* Cuanto más específico sea el *prompt,* mejores

serán los resultados. Sería más adecuado y eficaz, entonces, pedir: "escribe un ensayo sobre un tema X que aborde los siguientes subtemas S1, S2, S3, y que esté desarrollado en 20 páginas".

Entonces:

- Defina claramente el objetivo del *prompt*.
- Proporcione detalles relevantes y específicos.
- Incluya ejemplos si es necesario.

e) Creatividad

No hay que tener miedo de ser creativo con los *prompts*. Se pueden utilizar metáforas, analogías y juegos de palabras para obtener resultados más interesantes.

Así:

- Explore diferentes formatos de *prompts*.
- Experimente con diferentes estilos de lenguaje.
- Incorpore humor o ingenio si es apropiado.

f) Condiciones de formato

- Tono y estilo: el tono y estilo del *prompt* deben ser adecuados al contexto y al objetivo que se quiere lograr.
- Formato: el formato del *prompt* puede variar dependiendo del modelo de IA que se va a utilizar.
- Longitud: la longitud del *prompt* también puede afectar a los resultados.

g) Pruebas y ajustes

Es importante probar diferentes *prompts* y ajustar los parámetros del modelo para obtener los mejores resultados. Esto quiere decir que en la elaboración de *prompts* deben tenerse en cuenta cuestiones contextuales, culturales y coyunturales y que, además, cada *prompt* deberá ser flexible y dúctil para ajustarse una y otra vez hasta llegar a la obtención del resultado esperado.

> Un *prompt* efectivo es la clave para desatar el mayor potencial de la IA y obtener los resultados deseados.

De esta forma:

- Adapte el *prompt* al contexto y la tarea.
- Considere el conocimiento y las habilidades del modelo de IA.
- Use un tono y estilo apropiado para la audiencia.

h) Ética y responsabilidad

Como ya venimos explicando a lo largo de este libro, la herramienta no es el problema sino las posibilidades de uso que la herramienta habilita. Esto quiere decir que es posible lograr *prompts* que se ajusten plenamente a todas las premisas anteriores pero cuya intencionalidad influye para lograr contenido engañoso, malicioso, falso, discriminatorio o, directamente ofensivo. Por eso es importante utilizar los *prompts* de forma ética y responsable. ¿Cuál es la finalidad con la que, eventualmente, será utilizada esa información o ese contenido logrado? Nadie puede saberlo hasta el momento. Sí se puede aplicar la ética a la utilización de *prompts* y, es más, debe hacerse.

La ingeniería de los *prompts*

La mejor definición de *prompt* en IA es la que los considera como una herramienta fundamental para la interacción entre humanos e IA. Esto quiere decir que un *prompt* efectivo es la clave para desatar el mayor potencial de la IA y obtener los resultados deseados.

En este sentido, diseñar *prompts* efectivos es una habilidad que requiere práctica y experiencia. Se ha acuñado el término "ingeniería de *prompts*" para referirse a esta disciplina.

En resumen, un *prompt* en IA es mucho más que una simple instrucción. Es un poderoso instrumento que abre la puerta al potencial de la IA (o puede no lograr hacerlo, si no está bien diseñado) y de esta forma incidir en el futuro, en todas las áreas que afectan a los seres humanos.

Los *prompts* son una parte crucial para la interacción con modelos de IA, especialmente aquellos que se basan en lenguaje natural. Un *prompt* bien diseñado puede guiar al modelo hacia la respuesta

deseada, mientras que un *prompt* mal diseñado puede generar resultados inesperados o incluso irrelevantes.

Estructuras básicas de los *prompts*

a) Instrucciones:

Son la forma más simple de *prompt* y consisten en una orden directa al modelo. Ejemplo: "Escribe un ensayo sobre la literatura española del Siglo de Oro".

b) Preguntas:

Permiten al usuario obtener información específica del modelo. Ejemplo: "¿Cuál es la clasificación de las ciencias?"

c) Descripciones:

Proporcionan al modelo información contextual para que pueda generar una respuesta más precisa. Ejemplo: "Escribe un diálogo entre dos personajes que están discutiendo sobre el cambio climático".

Arquitecturas de *prompts*

Las estructuras básicas se pueden combinar para crear arquitecturas de *prompts* más complejas. Algunas de las arquitecturas más comunes son:

a) Secuencias

Son una serie de *prompts* que se ejecutan uno tras otro. Ejemplo: "Primero, escribe un resumen del informe. Luego escribe una crítica sobre él".

b) Jerarquías

Es la división de *prompts* en *sub-prompts* que se anidan unos dentro de otros. Ejemplo: "Escribe un artículo sobre la historia de la Revolución Francesa. En el primer párrafo, describe el contexto histórico. En el segundo párrafo, describe los eventos desatados a partir de 1789".

c) Plantillas

Proporcionan un marco predefinido para la respuesta del modelo. Ejemplo: "Escribe una carta

formal a la empresa X para realizar un reclamo. En la carta, debes incluir tus datos, tu caso específico y tus expectativas de resarcimiento".

d) Ejemplos de *prompts* para diferentes tareas

- Traducción:
  - Traduce este texto del español al inglés.
  - Traduce este artículo de noticias del inglés al francés.

- Escritura creativa:
  - Escribe un poema sobre la muerte y la vida que tenga formato endecasilábico y rima ABBA.
  - Escribe un guion para un cortometraje de comedia.

- Resumen de información:
  - Resume este artículo científico en 100 palabras.
  - Resume este video de YouTube en 3 puntos clave.

- Análisis de datos:
    - Encuentra las frecuencias en este conjunto de datos de estadísticas médicas.
    - Identifica los factores que influyen en la tasa de embarazo en la adolescencia.

- Generación de código:
    - Escribe un programa en Python que imprima los números del 1 al 100.
    - Crea una página web simple con HTML y CSS.

- Generación de un diálogo:
    - Escribe un diálogo entre dos personajes que están discutiendo sobre el cambio climático. Uno de los personajes cree que el cambio climático es una amenaza real, mientras que el otro personaje no cree que sea un problema grave. El diálogo debe ser de al menos 500 palabras.

- Escritura de un artículo:
    - Escribe un artículo sobre la historia de la inteligencia artificial. En el primer párrafo, describe los inicios de la IA, incluyendo los primeros modelos de IA y las aplicaciones que se les dieron. En el segundo párrafo, describe los avances recientes en IA, como el aprendizaje profundo y las redes neuronales artificiales. En el tercer párrafo, describe los desafíos que enfrenta la IA en el futuro.

La integración del Modelo Neuro-Tecno-Pedagógico en la Inteligencia Artificial: Un camino hacia la educación del futuro

En un mundo en constante transformación, en el que la tecnología y la información se erigen como pilares fundamentales del desarrollo, la educación se enfrenta al

desafío de adaptarse a las nuevas necesidades y demandas. En este contexto, la integración de la Inteligencia Artificial (IA) en los procesos educativos emerge como una oportunidad para revolucionar la enseñanza, el aprendizaje y la evaluación. Sin embargo, para aprovechar al máximo el potencial de la IA en la educación, es necesario hacerlo de manera responsable y efectiva. Es aquí donde nuestro Modelo Neuro-Tecno-Pedagógico cobra vital importancia.

> Para aprovechar al máximo el potencial de la IA en la educación, es necesario hacerlo de manera responsable y efectiva. Es aquí donde nuestro Modelo Neuro-Tecno-Pedagógico cobra vital importancia.

Este modelo está basado en tres principios fundamentales (Disparadores cognitivos, Procesos cognitivos y Tipos de pensamiento) y ofrece un marco teórico y práctico sólido para la integración de la IA en el ámbito educativo.

A continuación, analizaremos en profundidad las características del Modelo Neuro-Tecno-Pedagógico y su potencial para la innovación educativa. Abordaremos los beneficios de su aplicación en la IA, explorando

ejemplos concretos de cómo esta integración puede mejorar la experiencia de aprendizaje y potenciar el desarrollo de habilidades esenciales en los estudiantes.

a) Disparadores cognitivos: la chispa de la curiosidad

El Modelo Neuro-Tecno-Pedagógico propone el uso de preguntas como ¿qué?, ¿por qué? y ¿para qué?, ¿dónde?, ¿cuándo?, ¿quién/es? entre otras, las cuales son utilizadas para vehiculizar la intencionalidad de todo el proceso educativo, y además son implementadas como herramientas para estimular la curiosidad, la reflexión crítica y la búsqueda de conocimiento. En el contexto de la IA, estos disparadores cognitivos pueden ser integrados en plataformas educativas inteligentes para personalizar el aprendizaje, ofrecer contenido relevante a los intereses de cada estudiante y fomentar la participación activa en el proceso educativo.

También podríamos pensar, dentro del tema de los disparadores cognitivos, una taxonomía ampliada que va más allá de la Taxonomía de Bloom, expandiendo

la gama de preguntas y desafíos cognitivos para estimular la metacognición y la autorregulación del aprendizaje. Se incluyen niveles como:

- **Análisis metacognitivo**: ¿Cómo aprendo mejor? ¿Qué estrategias puedo utilizar para mejorar mi aprendizaje?
- **Evaluación crítica**: ¿Qué tan confiable es esta información? ¿Cómo puedo evaluar la validez de diferentes fuentes?
- **Creatividad y pensamiento divergente**: ¿Qué otras soluciones existen para este problema? ¿Cómo puedo generar ideas nuevas y originales?
- **Toma de decisiones**: ¿Qué opción es la mejor en este caso? ¿Cómo puedo tomar decisiones responsables y fundamentadas?

Estos disparadores cognitivos, cuidadosamente seleccionados y secuenciados, incentivan la curiosidad, la reflexión crítica y la búsqueda activa del conocimiento. Promueven un aprendizaje más profundo y significativo, por el que el alumnado no solo

memoriza información, sino que la comprende, la analiza, la critica y la aplica en diferentes contextos.

b) Procesos Cognitivos: un viaje desde el conocimiento básico hasta la creación.

El modelo abarca una amplia gama de procesos cognitivos, desde la adquisición de conocimientos básicos hasta la capacidad de crear, evaluar y colaborar, entre tantos otros que describimos a continuación. La IA puede potenciar estos procesos de diversas maneras, por ejemplo, mediante la automatización de tareas repetitivas, la provisión de feedback personalizado y la creación de entornos de aprendizaje colaborativos —en sus dimensiones física, digital y virtual inmersiva— integrando un real neuroaprendizaje, mediante el uso de recursos y medios en diversos recorridos tecno pedagógicos.

El MNTP abarca un espectro completo de procesos cognitivos, desde la simple memorización hasta la creación y evaluación crítica de información. Se

destaca la atención a habilidades de orden superior como:

- Pensamiento crítico: analizar, interpretar y evaluar información de forma crítica y reflexiva.
- Resolución de problemas: identificar problemas, analizarlos y proponer soluciones creativas y eficientes.
- Toma de decisiones: evaluar diferentes opciones y tomar decisiones responsables y fundamentadas.
- Comunicación efectiva: expresar ideas y pensamientos de forma clara, precisa y concisa.
- Colaboración: trabajar en equipo de forma efectiva para alcanzar objetivos comunes.
- Creatividad e innovación: generar ideas nuevas y originales, y ponerlas en práctica de forma efectiva.

c) Tipos de pensamiento: fomentando el pensamiento crítico y creativo

El Modelo Neuro-Tecno-Pedagógico pone énfasis en el desarrollo de diferentes tipos de pensamiento, como la observación, el análisis, la argumentación, la formulación de preguntas y la discriminación de información. La IA puede ser utilizada para crear experiencias de aprendizaje que fomenten estas habilidades, promoviendo el pensamiento crítico y creativo en los estudiantes.

El Modelo Neuro-Tecno-Pedagógico pone énfasis en el desarrollo de diferentes tipos de pensamiento, y la IA puede ser utilizada para crear experiencias de aprendizaje que fomenten estas habilidades, promoviendo el pensamiento crítico y creativo en los estudiantes.

En definitiva, la integración del Modelo Neuro-Tecno-Pedagógico en la IA abre un camino hacia la educación del futuro. Esta integración tiene el potencial de transformar la enseñanza y el aprendizaje, creando experiencias personalizadas, relevantes y motivadoras

para los estudiantes, y preparándolos para afrontar los desafíos del mundo actual.

d) Beneficios de la integración con IA

- Aprendizaje personalizado: la IA, al incorporar el modelo, podría adaptar las experiencias de aprendizaje a las necesidades e intereses individuales de cada estudiante.
- Mayor motivación e interés: los disparadores y procesos cognitivos, junto con la variedad de tipos de pensamiento, harían del aprendizaje una experiencia más atractiva y dinámica.
- Desarrollo de habilidades para el siglo XXI: la IA, con este modelo, fomentaría el pensamiento crítico, la creatividad, la resolución de problemas y la comunicación efectiva, habilidades que son esenciales en la actualidad.
- Evaluación formativa continua: la IA podría analizar el progreso de cada estudiante en tiempo real, brindando retroalimentación personalizada y oportuna.

- Inclusión y equidad: el modelo permitiría a la IA adaptarse a diferentes estilos de aprendizaje y necesidades especiales, promoviendo la inclusión educativa.

Dicho de otro modo, en el contexto actual, en que la inteligencia artificial (IA) experimenta un desarrollo exponencial, se hace patente la necesidad de explorar su potencial para la innovación educativa. La propuesta del Modelo Neuro-Tecno-Pedagógico (MNTP) puede ofrecer un marco teórico sólido para integrar la IA en los procesos de enseñanza y aprendizaje de manera efectiva y responsable.

El MNTP busca desarrollar un conjunto de habilidades transversales que son esenciales para el éxito en la sociedad actual. Estas habilidades van más allá del ámbito académico y preparan a los estudiantes para afrontar los desafíos del mundo real, a través de diferentes destrezas que hoy son imprescindibles, como la diversidad de enfoques, el pensamiento creativo y crítico, el pensamiento analítico que busca descomponer los problemas complejos, el pensamiento

sistemático, organizativo, estructural, el pensamiento reflexivo y metarreflexivo, el pensamiento estratégico, con capacidad de planificar y ejecutar acciones, y, fundamentalmente, el pensamiento flexible.

La flexibilidad cognitiva es la capacidad de adaptarse a diferentes situaciones y desafíos. Por medio de ella el alumnado aprende a pensar de forma crítica, creativa y estratégica, pero siendo capaz de adaptarse a diferentes situaciones, lo que les permite desenvolverse con éxito en un mundo en constante cambio.

> El Modelo Neuro Tecno Pedagógico busca desarrollar un conjunto de habilidades transversales que son esenciales para el éxito en la sociedad actual. Al incorporarlo al uso de la IA con fines pedagógicos se pueden adaptar las experiencias de aprendizaje a cada estudiante.

El aprendizaje personalizado que sería posible en esta integración de la que hablamos implicaría la posibilidad de —también— una adaptación individualizada. En efecto, la IA, al incorporar el modelo, puede adaptar las experiencias de aprendizaje a las necesidades, intereses, estilos de aprendizaje y ritmo

individual de cada estudiante. Esto se traduce en un aprendizaje más efectivo, eficiente y motivador, en el que cada alumno progresa a su propio ritmo y alcanza su máximo potencial.

e) Elementos clave a utilizar en el diseño de *prompt* en base al modelo Neuro-Tecno-Pedagógico

- Disparadores cognitivos: ¿Qué?, ¿Por qué?, ¿Para qué?, ¿Cuáles?, ¿Dónde?, ¿Cuánto?, ¿Cómo?, ¿Con qué?, ¿De qué forma?, ¿Quién/es?, ¿En qué medida?
- Procesos cognitivos: Conocer – Aplicar – Analizar – Sintetizar – Evaluar – Crear – Recordar – Comparar – Valorar – Decidir – Diferenciar – Reconocer - Descubrir – Interpretar – Criticar – Distinguir – Clasificar – Seleccionar – Descubrir – Argumentar – Explicar – Participar – Colaborar.
- Tipos de pensamiento: Observar y describir lo observado - Explicar e interpretar - Razonar a través de las evidencias-Realizar conexiones entre diversos elementos y/o conceptos-Considerar los dis-

tintos puntos de vista en relación o en un tema o problema - Discriminar lo esencial de lo accesorio - Formular conclusiones-Interrogarse y ser capaz de interrogar-Distinguir entre argumentos y prejuicios.

Además, en cada *prompt* podremos plantear alguna de las siguientes tareas: procesos creativos, resolución de problemas, generación de patrones novedosos, toma de decisiones, percepción sensorial, razonamiento emocional y emisión de juicios basados en conocimientos.

Es importante destacar que los *prompts* para IA son una herramienta poderosa que permite fomentar el aprendizaje activo, comprensivo y significativo. Al proporcionar a los estudiantes tareas desafiantes y estimulantes por medio de la aplicación del Modelo Neuro-Tecno-Pedagógico, podemos ayudarlos a desarrollar su pensamiento crítico, su creatividad y sus habilidades de aprendizaje.

Interacción entre humanos e inteligencias artificiales

A continuación, brindamos variados ejemplos, configurados con distintos niveles de complejidad, los cuales pueden ser aplicados en forma directa a los espacios de diálogo que brindan las inteligencias artificiales. De esta forma comenzaremos distintos procesos de colaboración entre la inteligencia biológica–natural humana y la inteligencia artificial.

Es necesario aclarar, sin embargo, que estos son solo ejemplos que pueden ser adaptados a los propios intereses o materias de estudio. Asimismo, pueden combinarse diferentes disparadores, procesos y tipos de pensamientos para complejizar los *prompts.* Las posibilidades son infinitas y se logran a través de la experimentación y la creatividad, introduciendo diferentes formatos, preguntas y problemas a resolver

Lo más importante es que los *prompts* sean desafiantes e inviten a pensar de forma crítica y creativa.

La estructura básica de un *prompt* que utiliza las bases del Modelo NTP, debería contener al menos los siguientes elementos:

a) Destinatarios del ejercicio

b) Contexto

c) Consigna

A continuación, ofrecemos diversos ejemplos logrados en la convergencia del Modelo Neuro-Tecno-Pedagógico y la inteligencia artificial:

Ejemplo 1

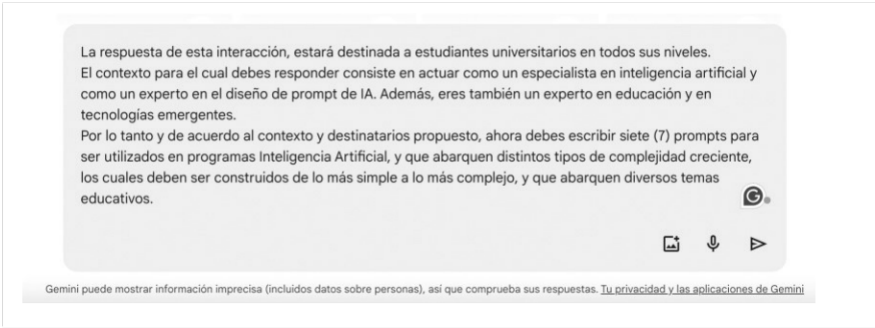
a) **Destinatarios del ejercicio**: "Estudiantes universitarios en todos sus niveles".

b) **Contexto**: "Actúa como un especialista en inteligencia artificial y diseño de *prompt,* pero, además, como un experto en educación y tecnologías emergentes".

c) **Consigna**: "De acuerdo al contexto propuesto, ahora debes escribir siete (7) *prompts* para ser utilizados en IA, y que abarquen tipos de

complejidad creciente, construidos de lo más simple a lo más complejo, y que a la vez abarquen diversos temas educativos".

Entonces nuestro *prompt,* en su forma más simple, se vería de la siguiente forma[25]:



En respuesta a este *prompt,* la IA Gemini ofrece 7 *prompts* de IA para estudiantes universitarios (de lo simple a lo complejo):

[25] Usando la herramienta de inteligencia artificial Gemini con sentido educativo y solo a los fines de un uso exclusivamente pedagógico, brindada por Google con todos sus derechos reservados de marca, la cual se obtiene del siguiente link: https://gemini.google.com/

1. Clasificación de imágenes: identifica las especies de aves en fotos.

Nivel: básico

Materia: Biología, Ecología

Descripción: el prompt busca que los estudiantes desarrollen un modelo de IA capaz de clasificar imágenes de aves en base a sus características. Esto implica entrenar el modelo con un conjunto de datos de imágenes de aves etiquetadas, y luego evaluarlo con un conjunto de imágenes nuevas.

2. Análisis de texto: resume artículos científicos.

Nivel: intermedio.

Materia: Ciencias en general.

Descripción: este prompt desafía a los estudiantes a crear un modelo de IA que pueda resumir artículos científicos de manera precisa y concisa. El modelo debe ser capaz de identificar las ideas principales del artículo, eliminar información irrelevante y generar un resumen coherente.

3. Generación de texto: crea historias creativas.

Nivel: Intermedio.

Materia: Lengua y literatura, Humanidades.

Descripción: en este prompt, los estudiantes entrenarán un modelo de IA para generar historias creativas a partir de indicaciones o frases simples. El modelo debe ser capaz de crear historias originales, coherentes y con una narrativa atractiva.

> Las posibilidades de la IA enriquecida por el MNTP son infinitas y se logran a través de la experimentación y la creatividad de los *prompts,* introduciendo diferentes formatos, preguntas y problemas a resolver.

4. Traducción automática: traduce textos entre idiomas.

Nivel: Intermedio.

Materia: Idiomas.

Descripción: el objetivo de este prompt es que los estudiantes desarrollen un modelo de IA para traducir textos de un idioma a otro. El modelo debe ser capaz de comprender el significado del texto original, transferirlo al idioma de destino y generar una traducción fluida y natural.

5. Asistente virtual: responde preguntas sobre un tema específico.

Nivel: Avanzado.

Materia: a elección del estudiante.

Descripción: en este prompt, los estudiantes crearán un asistente virtual de IA que pueda responder preguntas sobre un tema específico, como historia, química o literatura. El modelo debe ser entrenado con una gran cantidad de información sobre el tema y ser capaz de responder preguntas de manera precisa y completa.

6. Chatbot conversacional: simula una conversación con un humano.

Nivel: Avanzado.

Materia: Psicología, Ciencias de la Computación.

Descripción: el desafío de este prompt es que los estudiantes desarrollen un chatbot de IA que pueda simular una conversación natural con un humano. El chatbot debe ser capaz de comprender el lenguaje natural, responder de manera coherente y mantener una conversación interesante.

7. Detector de sesgos algorítmicos: identifica sesgos en conjuntos de datos.
Nivel: Avanzado.
Materia: Ética de la IA, Ciencias de la Computación.
Descripción: este prompt es de gran relevancia actual, ya que busca que los estudiantes creen un modelo de IA capaz de detectar sesgos algorítmicos en conjuntos de datos. El modelo debe ser capaz de identificar patrones que podrían generar discriminación o sesgos injustos en la toma de decisiones.

Estos *prompts* son solo ejemplos y pueden adaptarse a diferentes necesidades y niveles educativos. Es importante que los estudiantes reciban una guía adecuada durante el proceso de desarrollo de sus proyectos de IA, para que puedan comprender los principios éticos y las implicaciones sociales de estas tecnologías.

Ahora avancemos incrementando la complejidad en el diseño del *prompt* e incluyendo más elementos del Modelo Neuro-Tecno-Pedagógico.

Vamos a diseñar un *prompt* —de acuerdo con el esquema ya analizado de destinatarios-contexto-tipo de tarea— que le solicite respuesta a la IA[26] Gemini de la empresa Google, pero ahora con la consigna de incluir elementos del Modelo NTP (en esta parte de la consigna la idea es combinar disparadores, procesos cognitivos y tipos de pensamiento, según la intencionalidad educativa y pedagógica que el docente intente realizar con sus estudiantes).

Este nuevo *prompt,* en su forma más avanzada, se vería de la siguiente forma:

Ejemplo 2

Tema: Historia

Subtema: Revolución industrial

Destinatarios de la actividad: estudiantes universitarios de la carrera de Profesorado de Historia, los cuales se encuentran cursando el 3er. Año de su carrera.

[26] Cabe aclarar que el diseño de los *prompts* puede ser utilizado en cualquier herramienta de Inteligencia Artificial. Lo que va a variar es el tipo de respuesta, lo que se debe a las distintas arquitecturas y sus modelos de datos e informaciones, sus fechas de vigencia y su actualización.

Contexto: actúa como un especialista en historia y como un experto en el tema de la Revolución Industrial.

Tarea: resolución de problemas

Consigna a escribir en el *prompt* de IA:

> "Investiga cuáles fueron los principales factores que impulsaron la Revolución Industrial. Qué impacto tuvo en la sociedad y el medio ambiente. Cómo se compara la Revolución Industrial con otras revoluciones tecnológicas y evalúa su importancia histórica.
>
> Además, debes interpretar y explicar cómo fueron los factores globales del mundo en el período que se gesta la revolución industrial, pero ten en cuenta distinguir entre lo que pueden ser argumentos y prejuicios.
>
> Por último, ten en cuenta que la información – respuesta que brindes, será utilizada para resolver problemas de interpretación en relación a las bases históricas de una revolución".

Como podemos observar en este *prompt,* el mismo está conformado por diversos elementos que responden al Modelo Neuro-Tecno-Pedagógico. Asimismo, los podemos visualizar en el *prompt* por su identificación de colores que hemos utilizado para este ejemplo:

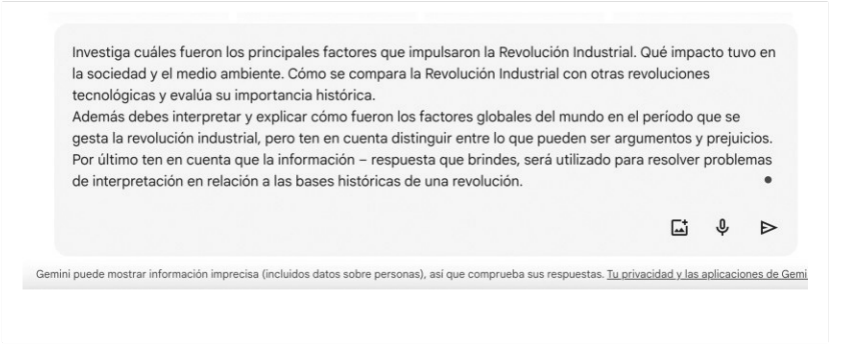
Disparadores cognitivos

Procesos cognitivos

Tipos de pensamiento

Tipos de actividad

Entonces, nuestro *prompt* cargado en el cuadro de diálogo de la IA Gemini, se vería de la siguiente forma:



Y la respuesta que nos brinda la inteligencia artificial Gemini se verá de la siguiente forma:

La Revolución Industrial: un motor de transformación

La Revolución Industrial, un período de profundo cambio socioeconómico y tecnológico, se gestó en Gran Bretaña a mediados del siglo XVIII y se extendió por el mundo a lo largo del siglo XIX. Este fenómeno marcó un antes y un después en la historia de la humanidad, impulsando un salto sin precedentes en la productividad, el comercio y la organización social.

Factores clave que impulsaron la Revolución Industrial:

- *Innovaciones tecnológicas: la invención de la máquina de vapor, la hiladora de algodón, el telar mecánico y otros dispositivos revolucionarios impulsaron la mecanización de la producción, reemplazando el trabajo manual y aumentando la eficiencia.*
- *Cambios agrarios: la Revolución Agrícola incrementó la productividad agrícola, liberando mano de obra para las nacientes industrias y generando excedentes para alimentar a las poblaciones urbanas en crecimiento.*
- *Capitalismo ascendente: el sistema capitalista, basado en la propiedad privada y la búsqueda del lucro, proporcionó el marco económico para la inversión en nuevas tecnologías y la expansión industrial.*
- *Abundancia de recursos naturales: Gran Bretaña poseía abundantes depósitos de carbón y hierro, materias primas esenciales para alimentar las máquinas de vapor y la industria siderúrgica.*

- *Espíritu emprendedor: la existencia de una clase media con iniciativa y capital dispuesto a invertir en nuevas empresas fue crucial para el desarrollo industrial.*

Impacto en la sociedad y el medio ambiente:

- *Urbanización: la industrialización atrajo a grandes masas de población del campo a las ciudades, generando un crecimiento urbano sin precedentes y dando lugar a nuevas formas de organización social.*
- *Condiciones laborales: las fábricas se caracterizaban por jornadas laborales largas, condiciones insalubres y bajos salarios, lo que provocó la explotación de trabajadores, especialmente mujeres y niños.*
- *Desigualdad social: la brecha entre ricos y pobres se amplió, con una concentración de riqueza en manos de los dueños de las fábricas y una precarización de las condiciones de vida de la clase trabajadora.*

> Un *prompt* diseñado con el Modelo Neuro-Tecno-Pedagógico debe contemplar: disparadores cognitivos, procesos cognitivos, tipos de pensamiento, tipos de actividad.

- *Degradación ambiental: la industrialización trajo consigo un aumento de la contaminación del aire y el agua, la deforestación y el agotamiento de recursos naturales, sentando las bases para los problemas ambientales que enfrentamos hoy en día.*

Comparación con otras revoluciones tecnológicas:

La Revolución Industrial se distingue de otras revoluciones tecnológicas por su magnitud y alcance global, transformando no solo los métodos de producción sino también la estructura social, económica y política de las sociedades. A diferencia de otras revoluciones, como la Revolución Neolítica o la Revolución Científica, la industrialización no se limitó a un área geográfica específica, sino que se extendió por todo el mundo, conectando economías y culturas a través del comercio y la migración.

Evaluación de su importancia histórica:

La Revolución Industrial es considerada uno de los eventos más importantes en la historia de la humanidad. Sus efectos se siguen sintiendo hoy en día,

desde la forma en que producimos bienes y servicios hasta la organización de nuestras ciudades y la estructura de nuestras sociedades.

La industrialización sentó las bases para el mundo moderno, impulsando el crecimiento económico, la innovación tecnológica y el desarrollo social.

Interpretación de factores globales:

Es importante analizar los factores globales del período previo a la Revolución Industrial, considerando tanto argumentos como posibles prejuicios. Algunos factores relevantes incluyen:

- *Expansión colonial: las potencias europeas, como Gran Bretaña, se beneficiaron de la explotación de sus colonias, accediendo a materias primas, mano de obra barata y nuevos mercados para sus productos manufacturados.*
- *Avances científicos: los descubrimientos científicos y las ideas de la Ilustración contribuyeron a crear un ambiente propicio para la innovación y el desarrollo tecnológico.*

- *Crecimiento del comercio internacional: el aumento del comercio global facilitó el intercambio de bienes, ideas y tecnologías, impulsando la difusión de la industrialización a otras regiones del mundo.*

Sin embargo, es importante reconocer que la Revolución Industrial también tuvo consecuencias negativas para muchas partes del mundo, como la explotación colonial, la degradación ambiental y la desigualdad social. Al analizar los factores globales, es crucial tener una perspectiva crítica y considerar diferentes puntos de vista.

Conclusión:

La Revolución Industrial fue un fenómeno complejo y multifacético que transformó profundamente la sociedad humana. Su impacto sigue resonando en la actualidad, y su estudio nos permite comprender mejor las bases del mundo moderno y los desafíos que enfrentamos como sociedad. Al analizar los factores que impulsaron la Revolución Industrial, su impacto y su comparación con otras revoluciones tecnológicas,

podemos obtener una perspectiva más completa de este evento histórico fundamental.

Por último, dejamos otros ejemplos de *prompt* que pueden ser utilizados según la intencionalidad educativa que se persiga, remarcando que se pueden obtener muchos beneficios educativos y de aprendizaje para los estudiantes si se combinan todos los elementos del Modelo Neuro-Tecno-Pedagógico, en conjunto con las potencialidades que hoy nos brindan las IA. También es necesario recordar que se debe continuar con las interacciones que surjan de las respuestas de las IA, para que los estudiantes puedan seguir aprendiendo a pensar, tomar decisiones, resolver problemas, emitir juicios de valor por medio de estas interacciones y, consecuentemente, formular nuevas preguntas a las IA.

Para optimizar los resultados, es recomendable centrar las prácticas educativas en función de aquello que las IA no pueden realizar, por lo menos hasta este momento en que estamos escribiendo el presente libro. La tarea que las IA no pueden realizar —todavía— es

PREGUNTARSE. Para preguntarse, todavía necesitan de humanos que hagan posibles estas interacciones.

Ejemplo 3

Destinatarios: diferentes niveles.

Contexto: *prompts* para IA en educación, de menor a mayor complejidad.

Consigna: realizar 4 *prompts.*

La tarea que las IA no pueden realizar —todavía — es preguntarse. Para preguntarse, todavía necesitan de humanos que hagan posibles estas interacciones.

Primer prompt:

Nivel inicial: observación y descripción.

Tema: Biología, la fotosíntesis.

Prompt: observa la imagen de una hoja bajo un microscopio. Describe las estructuras que puedes identificar y explica cómo se relacionan con la fotosíntesis.

Disparadores: ¿Qué observamos en la imagen de una hoja bajo un microscopio? ¿Qué estructuras podemos identificar? ¿Cómo se relacionan con la fotosíntesis?

Procesos: observar, describir lo observado, identificar, relacionar, realizar conexiones entre diversos elementos y/o conceptos.
Tarea: Percepción sensorial.

Segundo prompt:

Nivel avanzado: Razonamiento y argumentación.

Tema: Filosofía. La ética de la inteligencia artificial

Prompt: investiga los dilemas éticos que plantea el desarrollo de la inteligencia artificial. Elabora un argumento sólido sobre cómo podemos garantizar que la IA se use de manera responsable y beneficiosa para la humanidad. Propone un conjunto de regulaciones que consideres necesarias para controlar el desarrollo y uso de la IA.

Disparadores: ¿Qué dilemas éticos plantea el desarrollo de la inteligencia artificial? ¿Cómo podemos garantizar que la IA se use de manera responsable y beneficiosa para la humanidad? ¿Qué tipo de regulaciones son necesarias para controlar el desarrollo y uso de la IA?

Procesos: razonar, argumentar, evaluar, criticar.

Pensamiento: razonar a través de las evidencias, considerar los distintos puntos de vista en relación o en un tema o problema, formular conclusiones.
Tarea: toma de decisiones.

Tercer prompt:
Nivel creativo: generación de ideas.
Tema: Arte. Diseño de un videojuego educativo.
Prompt: diseña un videojuego educativo que ayude a los estudiantes a aprender sobre un tema específico de tu interés. Define el objetivo del juego, la mecánica de juego, los personajes, la narrativa y los elementos multimedia que se utilizarán.
Disparadores: ¿Qué tipo de videojuego educativo podría ayudar a los estudiantes a aprender sobre un tema específico? ¿Qué características debería tener para ser atractivo y efectivo? ¿Cómo se podrían integrar diferentes elementos multimedia para crear una experiencia de aprendizaje inmersiva?
Procesos: crear, imaginar, sintetizar, diseñar.
Pensamiento: generar patrones novedosos, discriminar lo esencial de lo accesorio.

Tarea: procesos creativos.

Cuarto prompt:

Nivel experto: investigación y desarrollo.

Tema: Ingeniería-Desarrollo de un robot para la limpieza de océanos

Prompt: investiga las tecnologías existentes para la limpieza de océanos y propone un diseño innovador para un robot que pueda realizar esta tarea de manera eficiente y respetuosa con el medio ambiente. Define los requisitos técnicos del robot, los materiales que se utilizarán para su construcción y el plan de desarrollo del proyecto

Disparadores: ¿Qué características debería tener un robot para la limpieza de océanos? ¿Qué tecnologías se podrían utilizar para su construcción? ¿Cómo se podría optimizar su diseño para que sea eficiente y respetuoso con el medio ambiente?

Procesos: Investigar, desarrollar, analizar, evaluar, proponer.

Pensamiento: Interrogarse y ser capaz de interrogar, formular conclusiones, argumentar y explicar.

Tarea: Resolución de problemas.

Ejemplo 4

Destinatarios: estudiantes universitarios

Consigna: 30 *prompts* de Nivel Básico, intermedio y avanzado.

Nivel básico:

I. Observación: ¿Qué observas en esta imagen del sistema solar? (Observar, Describir).

II. Comparación: ¿En qué se diferencian las células animales de las vegetales? (Comparar, Clasificar).

III. Explicación: ¿Por qué el cielo es azul? (Razonar, Explicar).

IV. Relación: ¿Cómo se relaciona la fotosíntesis con la respiración celular? (Conocer, Relacionar).

V. Definición: ¿Qué es la entropía y cómo se relaciona con el segundo principio de la termodinámica? (Conocer, Definir).

VI. Análisis: ¿Cuáles son las causas y consecuencias de la Revolución Industrial? (Analizar, Evaluar).

VII. Creatividad: Diseña un cartel publicitario para promover la lectura en tu universidad (Crear, Diseñar).

VIII. Toma de decisiones: ¿Qué carrera universitaria elegirías en base a tus intereses y habilidades? (Decidir, Valorar).

IX. Emoción: ¿Cómo se refleja la tristeza en la música? (Percepción sensorial, Razonamiento emocional).

X. Solución de problemas: ¿Cómo se puede reducir la contaminación del aire en tu ciudad? (Resolver problemas, argumentar).

Nivel Intermedio:

I. Interpretación: ¿Qué mensaje transmite este poema y cómo lo hace? (Interpretar, Criticar).

II. Patrones: ¿Qué patrones encuentras en la historia de las civilizaciones antiguas? (Identificar patrones, Descubrir).

III. Hipótesis: ¿Qué hipótesis puedes formular sobre la evolución del lenguaje? (Formular hipótesis, Razonar).

IV. Experimentación: diseña un experimento para probar la eficacia de un nuevo método de aprendizaje (Experimentar, Evaluar).

V. Debate: ¿Qué argumentos a favor y en contra existen sobre la pena de muerte? (Debatir, Argumentar).

VI. Solución creativa: ¿Cómo se puede utilizar la tecnología para mejorar la calidad de vida de las personas mayores? (Solución creativa, Diseñar).

VII. Toma de decisiones: ¿Qué decisión tomarías en este dilema moral y por qué? (Toma de decisiones, Valorar).

VIII. Emoción: ¿Cómo se puede utilizar la inteligencia artificial para comprender mejor las emociones humanas? (Percepción sensorial, Razonamiento emocional).

IX. Análisis crítico: ¿Qué sesgos o prejuicios se pueden encontrar en este artículo de opinión? (Analizar, Criticar).

X. Simulación: crea una simulación por computadora para modelar el comportamiento de un sistema físico (Simular, Modelar).

Nivel Avanzado:

I. Investigación: investiga y presenta un informe sobre los últimos avances en la investigación del cáncer (Investigar, Informar).

II. Análisis de datos: analiza un conjunto de datos sobre el cambio climático y extrae conclusiones (Analizar datos, Interpretar).

III. Modelado: crea un modelo matemático para explicar el comportamiento del mercado bursátil (Modelar, Simular).

IV. Diseño: diseña un robot que pueda realizar tareas domésticas de forma autónoma (Diseñar, Programar).

V. Emprendimiento: desarrolla un plan de negocios para una empresa que utilice inteligencia artificial (Emprender, Planificar).

VI. Solución creativa: ¿Cómo se puede utilizar la IA para combatir la pobreza en el mundo? (Solución creativa, Argumentar).

VII. Toma de decisiones: ¿Qué decisión tomarías en este escenario complejo y por qué? (Toma de decisiones, Evaluar).

VIII. Emoción: ¿Cómo se puede utilizar la IA para crear experiencias de aprendizaje personalizadas y emocionalmente atractivas? (Percepción sensorial, Diseñar).

IX. Tesis: escribe una tesis doctoral sobre un tema de investigación relacionado con la IA. (Investigar, Argumentar).

X. Debate: ¿Qué desafíos éticos presenta el desarrollo de la inteligencia artificial? (Debatir, Argumentar).

Ejemplo 5

En estos *prompts* se combinarán distintos grados de complejidad, con diferentes procesos cognitivos y tipos de pensamiento.

- *Prompt de observación y descripción (Nivel Básico)*

  Tarea: observa una imagen de un ecosistema natural (por ejemplo, un bosque, un arrecife de coral o una pradera) y describe detalladamente lo que ves.

  Disparadores cognitivos: ¿Qué?, ¿Dónde?

  Procesos cognitivos: observar, describir

  Tipo de pensamiento: observar y describir lo observado

- *Prompt de creatividad (Nivel Básico)*

  Tarea: imagina que eres un robot con emociones humanas. Escribe un breve relato sobre cómo navegas las complejidades de la inteligencia emocional.

Disparadores cognitivos: ¿Qué?, ¿Cómo?

Procesos cognitivos: crear, imaginar

Tipo de pensamiento: realizar conexiones entre elementos.

- *Prompt de razonamiento científico (Nivel Intermedio)*

  Tarea: explica el proceso de fotosíntesis en términos sencillos. ¿Por qué es fundamental para la vida en la Tierra?

  Disparadores cognitivos: ¿Por qué?, ¿Para qué?

  Procesos cognitivos: explicar, interpretar

  Tipo de pensamiento: explicar e interpretar

- *Prompt de resolución de problemas (Nivel Intermedio)*

  Tarea: diseña un algoritmo para resolver un problema matemático específico (por ejemplo, encontrar la raíz cuadrada de un número grande).

  Disparadores cognitivos: ¿Cómo?, ¿Con qué?

  Procesos cognitivos: analizar, crear

Tipo de Pensamiento: razonar a través de evidencias

- *Prompt de Ética en Inteligencia Artificial (Nivel Avanzado)*

  Tarea: discute las implicaciones éticas de la inteligencia artificial en roles de toma de decisiones. ¿Cómo deberíamos abordar la responsabilidad y la transparencia en los algoritmos?

  Disparadores cognitivos: ¿Por qué?, ¿En qué medida?

  Procesos cognitivos: argumentar, evaluar

  Tipo de pensamiento: Considerar distintos puntos de vista

- *Prompt de toma de decisiones (Nivel Avanzado)*

  Tarea: evalúa dos enfoques diferentes para abordar un problema ambiental global (como el cambio climático). ¿Cuál consideras más efectivo y por qué?

  Disparadores cognitivos: ¿Por qué?, ¿Cuáles?

Procesos cognitivos: valorar, Decidir

Tipo de pensamiento: distinguir entre argumentos y prejuicios

Análisis comparativo de la arquitectura y estructura de *prompts* en diferentes IA

Esta tabla no es exhaustiva y solo incluye algunas de las IA más populares del mercado.

> Los consejos básicos para diseñar *prompts* eficaces incluyen: investigar las capacidades del modelo de IA, probarlos y refinar los *prompts,* usar ejemplos, evitar el sesgo y, sobre todo, ser pacientes.

Asimismo, es necesario decir que la arquitectura y estructura de los *prompts* puede variar dentro de una misma IA, dependiendo de la tarea específica que se desea realizar.

Para que el diseño sea más riguroso, se aconseja consultar la documentación oficial de cada IA para obtener más información sobre cómo crear *prompts* específicos.

IA	Tipo de *prompt*	Estructura	Argumentos	Ejemplos
LaMDA (Google)	Conversa-cional	Fluido y abierto	Intención Entidades Contexto	"¿Qué tal si vamos al cine esta noche?"
GPT-3 (OpenAI)	Creativo	Libre y flexible	Tema Estilo Tono Longitud	"Escribe un poema sobre la soledad en la ciudad."
DALL-E 2 (OpenAI)	Descriptivo	Detallado y preciso	Objeto Escena Estilo Atributos	"Genera una imagen de un paisaje montañoso con un lago cristalino y un pequeño pueblo al pie de la montaña."
Bard (Microsoft)	Multipro-pósito	Adaptable a diferentes tareas	Tipo de tarea Instrucciones Parámetros	"Resume este artículo de Wikipedia sobre la historia de la IA."
Bloom (Hugging Face)	Informativo	Claro y conciso	Tema Preguntas Nivel de detalle	"¿Cuáles son los principales beneficios de usar la IA en

				la atención médica?"
Copilot (GitHub)	Codificación	Instrucciones específicas	Función deseada Parámetros Estilo de código	"Escribe una función en Python que calcule el promedio de una lista de números."
Gemini (Google)	Multipropósito	Similar a Bard, pero con mayor enfoque en la fluidez y la creatividad	Tipo de tarea Instrucciones Parámetros Estilo	"Escribe un diálogo entre dos personajes sobre el futuro de la IA."
Perplexity AI (Perplexity AI)	Multipropósito	Basado en preguntas y respuestas	Pregunta Contexto Estilo	"¿Cuál es el significado de la vida?"

Consejos prácticos para el diseño de *prompts*

a) Investigue las capacidades del modelo de IA:

• Comprenda las fortalezas y debilidades del modelo.

• Identifique los tipos de *prompts* con los que funciona mejor.

• Consulte la documentación del modelo para obtener información específica.

b) Pruebe y refine sus *prompts:*

• Experimente con diferentes variaciones del *prompt.*

• Observe cómo responde el modelo a cada variación.

• Realice ajustes para mejorar la precisión y la creatividad.

c) Use ejemplos:

• Proporcione ejemplos concretos de lo que desea que la IA produzca.

• Cuanto más específicos sean los ejemplos, mejor será la comprensión del modelo.

• Use ejemplos relevantes para la tarea que se está realizando.

d) Evite el sesgo en los *prompts:*

• Sea neutral e imparcial en la redacción del *prompt.*

• Evite usar lenguaje discriminatorio o estereotipado.

• Promueva la equidad y la inclusión en sus *prompts.*

e) Sea paciente:

• El diseño de *prompts* efectivo requiere tiempo y práctica.

• No se desanime si no obtiene resultados perfectos al principio.

• Continúe aprendiendo y experimentando para mejorar sus habilidades.

f) Últimas recomendaciones generales

• Los *prompts* pueden adaptarse a diferentes niveles educativos y áreas de conocimiento.

• Es importante proporcionar a los estudiantes la información y los recursos necesarios para completar las tareas.

• La evaluación de los resultados debe basarse en el cumplimiento de los objetivos del *prompt,* la creatividad y la originalidad del trabajo realizado.

Referencias bibliográficas

de Elorza Feldborg, G. (2016). *Revolución del aprendizaje en tiempos de lo digital.* Ediciones Diapasón

de Elorza Feldborg, G. (2021). *Educación, Neurociencia y nuevas tecnologías.* Ediciones Diapasón

de Elorza Feldborg, G. (2023). *Metaversos y Educación: ¿cómo no perderse en los nuevos laberintos de la virtualidad inmersiva?* Ediciones Diapasón.

Capítulo 7

¿Creando Monstruos?
La paradoja de la Inteligencia Artificial como el nuevo Frankenstein digital

¿Era yo entonces un monstruo,
un error sobre la tierra, un ser del que todos
los hombres huían y a quien todos los hombres
rechazaban?
(…)
Usted busca conocimiento y sabiduría,
como lo busqué yo;
y espero de todo corazón
que el fruto de sus deseos no sea
una víbora que le muerda, como lo fue para mí.

Frankenstein. Mary Shelley.

Ya hemos expuesto —esperamos que con claridad— cómo es (en presente) el mundo que se avecina y cuán cerca lo tenemos. Hace años que venimos hablando acerca de la inserción de la tecnología en todos los ámbitos de la vida humana, con especial atención a lo que ella puede reportar —y, de hecho, reporta— a la educación. Desde ese mismo tiempo venimos considerando que una educación de espaldas a la tecnología —todas las tecnologías hasta la IA, y las

que vendrán y todavía no tenemos en el radar— está destinada al fracaso y a la irrelevancia. Particularmente en este libro hemos intentado mostrar los beneficios de nuestro Modelo NTP amalgamado con la IA y hemos aportado ejemplos tangibles de posibles intervenciones educativas. Junto a los beneficios, fuimos deslizando algunas señales de alerta. Por ello decidimos terminar este trabajo centrándonos en las señales de alerta, para que ellas sean una brújula, un marco ético desde el cual se orienten todas las aplicaciones que hagamos de la IA a la educación en particular y a la vida en general.

En este libro hemos intentado mostrar los beneficios de nuestro Modelo NTP amalgamado con la IA y hemos aportado ejemplos tangibles de posibles intervenciones educativas.

Mucho se habla y se ha hablado de la posibilidad de que las antiguamente llamadas "nuevas tecnologías" lleguen a un desarrollo tal que en algún punto del derrotero escapen a las propias previsiones de sus creadores. Para ello, y a falta de posibilidades ciertas de vaticinios futuros, se ha

echado mano de diversas metáforas e imágenes tomadas de la literatura y el mito que, en definitiva, trasuntan el inconsciente colectivo, el pensamiento arquetípico, los temores más atávicos y las esperanzas más fantásticas.

La novela de Mary Shelley, Frankenstein, es otro de los ejemplos que suelen ponerse cada vez que el accionar humano en nuevos descubrimientos y tecnologías es puesto bajo la lente del microscopio como consecuencia de los temores que todo lo nuevo suscita, por desconocido, pero fundamentalmente por las posibilidades de trazar líneas imaginarias que proyecten a futuro —más lejano o más cercano— el estado actual de la tecnología y, específicamente, de la Inteligencia Artificial.

Frankenstein o el moderno Prometeo es un relato dentro de un relato, en el que Victor Frankenstein, exhausto luego de los acontecimientos, es rescatado por una embarcación y decide contarle su historia de infortunios al capitán Robert Walton.

Victor, sabemos por su propio relato, es un joven conspicuo, interesado en la ciencia, la magia y la

filosofía, en una época en que las relaciones entre ellas no estaban del todo delimitadas. Sin embargo, es interesante conocer sus recuerdos de cuando comienzan sus estudios y él expone a sus maestros qué lecturas realizaba. A estos comentarios sus nuevos maestros responden que no esperaban encontrarse a un discípulo de Paracelso en plena época de la Ilustración y de la ciencia. Es que la novela tiene un trasfondo histórico de término de una época y comienzo de otra: el ascenso de la razón como medida de todas las cosas. Y este no es un dato menor.

El *leit motiv* de Victor era descubrir el misterio de la vida, encontrar dónde residía en cada ser humano el principio de la vida, y así comienzan sus investigaciones.

La historia es conocida: *puzzle* de restos humanos animados con corriente eléctrica, Victor por fin lograba su cometido, aunque el producto final pronto se le escapa de las manos, con toda la independencia que da la vida y la consciencia.

La paradoja de esta criatura cuasi humana es que puede pensar e incluso sentir y, por ello, puede confrontar con su creador, lastimado por la realidad de ser

quién es y de tener los límites que tiene. La creación increpa a su creador para que se haga cargo de su responsabilidad. En ese instante Victor habría recordado sus sentimientos a medida que hacía los descubrimientos: "*Cuando me encontré con un poder tan asombroso en las manos, durante mucho tiempo dudé sobre cuál podría ser el modo de utilizarlo".*

> Decidimos terminar este trabajo centrándonos en las señales de alerta, para que ellas sean una brújula, un marco ético desde el cual se orienten todas las aplicaciones que hagamos de la IA a la educación en particular y a la vida en general.

Los diálogos entre la criatura y su hacedor marcan un punto crucial en la historia: Frankenstein se enfrenta a la realidad de las consecuencias de sus acciones y se debate entre cumplir la demanda del monstruo y enfrentar las implicaciones morales de crear otro ser como él. La novela es mucho más que una historia fantástica, plantea cuestiones profundas sobre la responsabilidad, la ética y las derivaciones de jugar a ser Dios.

¿Y por qué nos interesa esta historia en un libro sobre Inteligencia Artificial?

Nick Bostrom (2014) explica el desarrollo de la IA —la superinteligencia— en términos de un pasaje, una trasposición de un umbral en un solo sentido, sin posibilidad de retroceso. Como si la tecnología fuera sacando "bolillas" de un bolillero, al azar: hasta ahora, reflexiona, siempre hemos sacado bolillas blancas, pero puede ser que al pasar por el portal saquemos una negra, y ya no haya vuelta atrás. Nada será igual entonces. Por supuesto que el filósofo sueco no aboga por un cese de las investigaciones o una vuelta al pasado. Todo lo contrario, pero sí advierte sobre lo que él denomina "riesgo existencial" (2002), es decir, "*aquel en el que un resultado adverso aniquilaría la vida inteligente originada en la Tierra o reduciría permanente y drásticamente su potencial".*

El monstruo creado por Shelley representa una creación artificial que posee un grado de conciencia, y por eso puede reclamar a su creador un lugar en la existencia. Análogamente podemos pensar que la IA potencialmente podría desarrollar una nueva forma de conciencia sintética, que le permita entre muchas cosas saber cuál es su lugar y conocer las formas de

convivencia con las personas. Asimismo, podría buscar en un futuro no muy lejano ser reconocida y comprendida por los seres humanos, a pesar de ser diferente en su naturaleza y funcionamiento.

Fue Raymond Kurzweil quien, a principios del siglo XXI auguró la posibilidad de creación de máquinas inteligentes capaces de autoreproducirse, en una espiral de "superinteligencia" —o IA fuerte— que, en teoría, podría prescindir del ser humano para su existencia.

Una creación argumentando con su creador, que ese es el escenario planteado en la novela fantástica, nos pone frente a la paradoja de una creación independiente e incluso superior. Por eso es importante la discusión acerca de la IA, que a veces puede ser malentendida o vista con temor debido a su capacidad para superar a los humanos en ciertas tareas o su aparente amenaza para ciertos empleos. La realidad es que el desarrollo de la IA no es un tema de moda, sino de modos en que este tipo de tecnología emergente pasa a formar parte de un nuevo laberinto cibernético de espejos sociales del siglo XXI. Y el nuevo sistema operativo social, el del individualismo en red, ya no se organiza en pequeñas

células sino capilarmente, de manera globalizada e hiperconectada, a través de internet, las redes sociales, los dispositivos móviles y los nuevos territorios inmersivos como el Metaverso

Haciendo un parangón, entonces, entre la literatura y la realidad, diremos que los creadores y desarrolladores de las tecnologías en general, y de la IA en particular, también tienen la responsabilidad de asegurarse de que ella se utilice de manera ética y responsable, considerando el impacto que puede tener en la sociedad y en su configuración. Pero no solamente en su uso, sino también en el propio desarrollo, es decir, en los objetivos que se tienen en la mira al trabajar en esos programas y en la previsión de posibles variables no deseadas.

En resumen, el diálogo entre el monstruo y su creador nos enfrenta, lateralmente, a la responsabilidad, la comprensión y las implicaciones éticas en la creación y el uso de la inteligencia artificial, de seres artificiales o de vida sintética. Aunque Frankenstein es una novela clásica escrita por Mary Shelley en el siglo XIX, su temática sigue siendo relevante en el contexto actual de la IA.

Al igual que Victor Frankenstein, los científicos e ingenieros de la IA crean sistemas y programas que pueden realizar tareas complejas y aprender de forma autónoma. Estos sistemas de IA pueden tomar decisiones, procesar información y realizar tareas específicas de manera similar a los seres humanos.

Sin embargo, al igual que el monstruo de Frankenstein, la IA también plantea cuestiones filosóficas:

1. Responsabilidad: ¿Quién es responsable de las acciones de la IA? Tal como Frankenstein, los creadores de la IA pueden enfrentarse a consecuencias no esperadas de sus creaciones. Si se asume ese riesgo en virtud de los innumerables beneficios, ¿quién asume el posible daño? ¿está previsto? ¿existe un plan de contingencia? Si las previsiones acerca de una superinteligencia prescindente del ser humano están a la puerta, ¿quién (y cómo) responde frente a sus acciones independientes?

2. Impacto social: la IA tiene el potencial de cambiar la forma en que vivimos y trabajamos. Por eso la IA puede ser vista como una fuerza disruptiva que amenaza la estabilidad

social y económica. ¿Lo es? ¿En qué forma? En todo caso, es evidente que vamos hacia nuevas configuraciones sociales mediadas por la tecnología. Prever los aspectos negativos no es hacer profecías apocalípticas sino estar preparados para morigerarlas. En lenguaje de Bostrom, anticiparse a la posibilidad de la bolilla negra antes de trasponer el umbral.

> La realidad es que el desarrollo de la IA no es un tema de moda, sino de modos en que este tipo de tecnología emergente pasa a formar parte de un nuevo laberinto cibernético de espejos sociales del siglo XXI.

3. Ética en la creación de la IA: los científicos e ingenieros de la IA deben considerar los dilemas éticos y morales asociados con la creación de seres artificiales. Esto incluye la consideración de la conciencia, la privacidad, la autonomía y el propósito de la IA.

4. Sesgos y discriminación: la IA, hasta el momento, es entrenada por personas que, como tales, tienen sesgos y prejuicios. Esto permite pensar que la IA puede reproducir esos sesgos de manera discriminatoria de acuerdo a los datos con que es conformada. El planteo moral aquí tiene que ver con las decisiones

automatizadas que, reproduciendo el sesgo, impliquen discriminación.

5. Privacidad y seguridad: ¿Cómo se gestiona la privacidad y la seguridad en un ecosistema en el que las máquinas —posiblemente independientes— manejan grandes cantidades de datos personales e información?

6. Automatización y futuro del empleo: la tecnología suplanta la antigua cadena de montaje del capitalismo moderno. Ya no es posible pensar el trabajo en esos términos, puesto que la automatización de tareas repetitivas reemplaza mano de obra humana, incluso mano de obra calificada. Es de esperar que la IA produzca desplazamientos laborales significativos. ¿Qué estrategias se están pensando para reestructurar el mercado laboral? Es más, ¿qué estrategias se están pensando para adecuar la educación a estos nuevos escenarios? ¿Se están pensando?

7. Decisiones autónomas: si la IA puede tomar decisiones autónomas ¿cómo decidirá en situaciones éticamente complejas? ¿Existe la posibilidad de programar esos sistemas para que respondan a esas situaciones éticamente complejas? ¿Bajo qué

paradigma? ¿El utilitarista? ¿El liberal? ¿El de Mill? ¿El de Rawls? Se podrían citar muchos otros, por supuesto. Si la ética aplicada es una disciplina eminentemente deliberativa (Gracia, 2009), en la que se sopesan posibles decisiones y cursos de acción frente a un hecho: ¿qué variables podrá tomar en cuenta la IA en ese proceso de toma de decisiones? ¿Cuál sería la respuesta propia y adecuada?

8. Transparencia y explicabilidad: los sistemas de IA en general funcionan como "cajas negras" en cuanto al manejo de datos y algoritmos y, por consiguiente, en términos de cómo llegan a sus decisiones. Los usuarios estamos expuestos —como en el panóptico foucaultiano— a una constante sociedad de la vigilancia sustentada en una profunda asimetría: quien posee los "datos" y observa, y quien es observado. La capacidad de explicar y comprender el razonamiento detrás de las decisiones de la IA, pero incluso el manejo de esos datos y algoritmos, es importante para garantizar la confianza y la rendición de cuentas. Existen ya muchos teóricos advirtiendo sobre las consecuencias de la opacidad de la IA: Harari (2017) lo

llama "dataísmo" o "religión de los datos", en el sentido de reducir toda la realidad a un flujo de datos. Danaher (2016) por su parte, habla de "algocracia" y pone el acento en que los sistemas algocráticos tienen muchos beneficios, pero, y por lo mismo, también tienen consecuencias negativas que no siempre están a la vista. De ahí la necesidad de transparencia y explicabilidad que den sustento a una "ética de datos" (Morte Ferrer, Monasterio Astobiza, Ausín, 2017). Retomando el concepto de "caja negra" utilizado al principio, y a fin de poder someter al escrutinio ético los comportamientos robóticos, sería necesario que exista una suerte de registro de actuaciones y de procedimientos lógicos involucrados en las acciones, a fin de fijar responsabilidades, pero también a fin de corregir posibles errores, fallos o consecuencias negativas.

9. Roboticización de las interacciones humanas: la comunicación humano/ máquina (CHM) ya excedió los límites del diseño, implementación, desarrollo y evaluación de sistemas informáticos interactivos. En la actualidad estamos sumergidos en otro paradigma totalmente diferente, en el cual la interacción está en el centro de la escena. A me-

dida que la IA avanza, surge la cuestión de si la sustitución de interacciones humanas por interacciones con máquinas puede tener un impacto negativo en la sociedad y las relaciones humanas. La IA, por cierto, en tanto que "tecnología disruptiva" incide y transforma los modos de ser y hacer de los seres humanos, y esto se extiende hasta el ámbito de las relaciones e interacciones. Al punto de que ya mismo se está hablando, en el ámbito del Derecho, de "personas electrónicas" (Valente, 2019) como sujetos de derechos y responsabilidades. En suma, la transformación de las relaciones interhumanas es un hecho irrefutable y, sin embargo, nos resta comprender —porque en la actualidad carecemos de la perspectiva que da el tiempo transcurrido— si estos cambios sociales son beneficiosos o no o si los aspectos positivos y negativos se dan equilibradamente. ¿Cómo ya cambió y habrá de cambiar todavía la sociedad en tiempos de IA? ¿Cómo se han transformado las redes sociales (en el sentido de entramados de contención social) en esta era del "robo-sapiens"? ¿Qué impacto tiene en las subjetividades? ¿Cómo afecta la salud mental? ¿Qué clase de alianzas, agrupaciones, célu-

las conformarán una nueva "colmena" social a partir de la interacción con IA? ¿Cómo serán las nuevas configuraciones familiares? ¿Cómo la educación? ¿Cómo la escuela? ¿Qué del amor y la amistad? Las preguntas son muchas y todavía no pueden ser contestadas. Sin embargo, habrá que preguntarse con espíritu crítico, sabiendo de antemano que ya nada será igual a partir de estos desarrollos. El filósofo Byung Chul Han (2014) llama a las nuevas sociedades "enjambre digital", destacando que lo que tenemos ahora es una suma de individualidades que no llegan nunca a conformar un "nosotros" y que, además, está sometida a una hipercomunicación constante y sin pausa que no habilita el pensamiento y la voz y, por tanto, impide el pensamiento crítico y la contrahegemonía.

> Sin embargo, al igual que el monstruo de Frankenstein, la IA también plantea cuestiones filosóficas sobre la responsabilidad, el impacto social, la ética en la creación de la IA, los sesgos y discriminación, la privacidad y seguridad, la automatización y futuro del empleo, las decisiones autónomas, la transparencia y explicabilidad, la roboticización de las interacciones humanas.

10. Equidad y acceso: ¿Es la IA una tecnología democrática? Es decir, ¿es posible pensar que la IA estará al alcance de todas las personas, de todas las edades, de todos los rincones más lejanos de este mundo? Si no es así, y tenemos todas las sospechas de que así no será, porque estamos en un sistema capitalista cada vez más desigual, es lícito preguntarnos ¿quiénes tendrán acceso y quiénes no? ¿cómo la utilizarán aquellos que sí tendrán acceso, porque revisten una situación social, económica o geográfica aventajada? ¿qué lugar en ese mapeo ocuparán los que no tendrán acceso? ¿qué determinará el acceso o no? Por otra parte, existe otra asimetría quizás poco tenida en cuenta: ¿quién tiene acceso al desarrollo de la IA? ¿qué países? ¿qué conglomerados económicos? En este punto, el panorama parece distópico. Nada hace pensar que la inequidad pueda solucionarse, como no se solucionó —sino que se profundizó— con la globalización. Así y todo, las acciones deben ir encaminadas a democratizar el acceso y a solucionar la inequidad. Evitar ampliar las brechas es un objetivo de mínima,

pero no puede convertirse en el horizonte de expectativa.

11. Ética en la investigación y desarrollo de IA: en la historia del siglo XX hubo algunos hitos que dieron lugar a reacciones de la comunidad científica respecto de las metodologías de investigación. En el juicio de Nüremberg, por ejemplo, se ventilaron las atrocidades que el régimen nazi realizaba sobre los prisioneros, amparado en la supuesta "investigación científica".

> También existen otros temas relacionados con la IA en debate filosófico: equidad y acceso, ética en la investigación y desarrollo de IA, manipulación de la información, autonomía y toma de decisiones, diversidad cultural y límites éticos, identidad y conciencia, responsabilidad en sistemas autónomos, dependencia y deshumanización, control y poder, impacto medioambiental, intimidad y vigilancia masiva, impacto en el trabajo humano, equidad en la distribución de beneficios.

Durante cuarenta años, del 30 al 70 del siglo XX, tuvo lugar el llamado "Experimento de Tuskegee", en el que se mantuvo a una población afroamericana que padecía sífilis sin medicación —solo con placebos y por supuesto sin saberlo ellos— a fin de estudiar la evolución

natural de la enfermedad. Estos hechos, y otros que se pueden citar, pusieron a la comunidad científica a pensar un marco legal adecuado que preservara la integridad humana y mantuviera los afanes investigativos dentro de ciertos límites tolerables. De allí surgen algunas premisas para la investigación científica que cristalizaron en el llamado "Principialismo" de Beauchamp y Childress (1999). Estos principios, que también se pueden aplicar a la investigación y desarrollo de la IA son: No maleficiencia (no hacer mal), Beneficencia (hacer bien), Justicia (equidad en el acceso), Autonomía (respeto a las decisiones de las personas). Se pueden complementar, asimismo, con las llamadas "Leyes de Asimov" sobre la robótica. Curiosamente, Isaac Asimov es un escritor de Ciencia Ficción que instala estas leyes en dos de sus relatos. Uno es "Círculo vicioso", de 1942 y el otro es "Los robots y el imperio", de 1985. Estas leyes son:

a) Un robot no debe dañar a la humanidad o por omisión, permitir que la humanidad sufra daños.

b) Un robot no puede perjudicar a un ser humano o, por inacción, permitir que algún humano sea lastimado.

c) Un robot debe obedecer órdenes dadas por un humano, excepto cuando estas vayan en contra de la ley.

d) Un robot debe proteger su propia existencia solo cuando esa protección no entre en colisión con la ley.

12. Manipulación de la información: la IA puede ser utilizada —y de hecho ya lo es— para manipular y generar información falsa o engañosa, lo que plantea preocupaciones sobre la confiabilidad y la integridad de la información en la sociedad. En este punto cobran nuevo vigor las diferentes teorías de análisis que vieron la luz con el desarrollo de los medios masivos de comunicación. Veamos solo algunas:

a) Teoría de la aguja hipodérmica de Harold Lasswell: los medios emiten estímulos que son recibidos —incluso sin ser advertidos— por una masa compacta y acrítica que responde a ellos de manera uniforme. Son como pequeñas dosis inyectadas por una aguja hipodérmica.

b) Determinismo tecnológico de Harold Innis: las tecnologías contribuyen a modificar o establecer estructuras económicas, sociales, políticas, etc., e influyen en la representación e interpretación del mundo.

c) "El medio es el mensaje", dirá Marshall McLuhann: el medio que logra la supremacía en una época será el que estructure las consciencias individuales, actuando de filtro en la percepción del mundo.

d) La sociedad del espectáculo de Guy Debord: los medios instalan un espectáculo, como una realidad paralela, que narcotiza toda reacción y es funcional a la lógica del mercado, del consumismo, del pensamiento en manada.

e) La "teoría del simulacro" de Jean Baudrillard: la realidad, dice, ha sido "asesinada". En su lugar se ha instalado un simulacro, una hiperrealidad que anula la posibilidad de distinción entre la verdadera realidad y el propio simulacro.

f) El periodismo como género propagandístico, de Chomsky: esa propaganda está cooptada por los grupos y corporaciones económicas mundiales, y busca

instalar una realidad, transmitir determinadas ideas y valores, inclinar la opinión pública hacia un lado u otro, etc.

g) Infocracia de Byung-Chul Hann: ya no estamos sometidos a un régimen disciplinario al estilo del panóptico de Foucault, sino sometidos a la tiranía de la información. La comunicación —y sus medios— son un dispositivo de vigilancia.

La IA ha demostrado capacidad de generar contenido falso muy convincente, como imágenes generadas por computadora o videos manipulados, y esto plantea desafíos para distinguir entre lo real y lo falso, y pone en entredicho los principios de confianza, autenticidad y transparencia en línea. Por otra parte, como en varios de los otros dilemas éticos, en este punto se podría rozar el delito, según el caso.

13. Autonomía y toma de decisiones: a medida que la IA se vuelve más sofisticada, y se avizora la creciente independencia que puede adquirir, se plantean preguntas sobre si las máquinas deben tener autonomía para tomar decisiones y en qué medida se debe permitir esta autonomía, toda vez que todas las

decisiones, directa o indirectamente, afectan a los seres humanos y a su medio.

14. Inteligencia Artificial, diversidad cultural y límites éticos: vivimos una cultura plural, un mundo plural, el emergente de muchas ideologías con necesidad y pretensión de ser reconocidas, escuchadas y respetadas. Sabemos, además, que la ética heterónoma ha ido cediendo paso frente a una ética autónoma o, mejor dicho, muchas éticas autónomas relativas, de acuerdo con las épocas, los grupos, las historias. Las diversidades y las minorías, étnicas, sexuales, sexogenéricas y sexoafectivas han dado una dura batalla para ser visibilizadas y reconocidas en sus derechos. El fenómeno de los feminismos no ha parado de crecer, obteniendo grandes avances tendientes a cerrar la brecha y terminar con las asimetrías de una sociedad desigual. En este contexto ¿cuál sería el consenso general para establecer los límites de una ética de la IA para todos y todas en todos los tiempos? Es más, con sistemas políticos y económicos que tienden cada vez más al individualismo y cada vez menos al bien común ¿cómo sería buscar el bien común en el

desarrollo y uso de la IA? Si cada grupo establece sus propias reglas, sus propias iniciativas, su propio objetivo y hoja de ruta ¿cómo delinear consensos generales? ¿cómo no reproducir sesgos? ¿cómo beneficiar a todos?

15. Identidad y conciencia: los partidarios de la IA fuerte creen que la actividad mental es de tipo computacional y, por tanto, los sentimientos, las emociones y la consciencia también lo son, de manera que sería eventualmente posible reproducirlos en una máquina inteligente. Los partidarios de la IA débil, por su parte, estiman que la consciencia no es reproducible, porque es un rasgo exclusivamente humano. Es posible, entonces, que existan sistemas de IA que emulen la consciencia humana y, frente a ello, la discusión que se impone es una de carácter antropológico: ¿dónde reside la humanidad? ¿Cuál es el rasgo distintivamente humano, sin el cual seríamos poco más que una máquina? Y, entonces, ¿qué pasaría en el mundo si ese rasgo —en este caso la consciencia— pudiera ser extra-humana? El argumento de Searle (1980) de la "Habitación china" parece ir en contra de esta presunción de la IA fuerte: supongamos que estamos encerrados en una habitación con una sola hendija que

comunica con el exterior, desde la que recibimos y a su vez enviamos fuera papeles con inscripciones en chino. En la habitación, además, hay instrucciones para poder leer y escribir en chino, y estas instrucciones están escritas en nuestra lengua materna. Quienes están fuera, argumenta Searle, no tienen razón para pensar que quien está adentro no sabe chino y no entiende el chino, porque sus preguntas son respondidas. Así funcionan las máquinas: pueden responder porque tienen instrucciones, pero ni entienden ni tienen una real interacción. Sin embargo, Searle sigue desarrollando esta teoría en escritos posteriores, configurando su propia filosofía de la mente, y llega a conclusiones sorprendentes: si la mente es solo un proceso biológico sujeto a leyes, es dable pensar, entonces, que es posible replicar esos procesos biológicos causales en las máquinas.

Pues bien, avancemos un poco más por este camino: ¿qué es la mente? ¿qué es la consciencia? ¿dónde reside? Con Bostrom podríamos señalar dos escenarios: uno, en el que el pasaje de la inteligencia humana a la superinteligencia es gradual, el cual daría lugar a un proceso de habituación entre ambas. El otro escenario

es el de la instantaneidad, y en ese caso, la posibilidad de errores y catástrofes es más plausible, por el escaso tiempo de que se dispondría para la preparación.

Ahora bien, si se da ese paso a la superinteligencia, hacia una máquina con identidad y consciencia, ¿podríamos estar ante el "giro traicionero" (Bostrom, 2016)? ¿Podría defender otros valores que aquellos que le transmitimos? ¿Podría generar conductas inesperadas, no aquellas para las cuales fue creada? ¿Podría suplantarnos? ¿Podría destruirnos?

16. Responsabilidad en sistemas autónomos: la IA ya está siendo aplicada en sistemas autónomos, como vehículos o robots. Evidentemente, es un beneficio y adelanto en la evitación de tareas repetitivas o desagradables, y en la comodidad, incluso la precisión, con que las tareas habituales pueden ser ejecutadas. Sin embargo, ¿quién es responsable en caso de accidentes o daños causados por estas máquinas? Se necesita legislación nueva para eventos nuevos, pero también se necesitará adecuar infraestructuras y servicios. No obstante, los desafíos que plantea la IA en este caso se relacionan más con cuestiones filosóficas de interacción

máquina-humano y sus eventuales consecuencias positivas y negativas.

17. Dependencia y deshumanización: ya no se trata solamente de las seis o siete horas que pasamos los seres humanos en promedio conectados a internet, sino de la cantidad de horas en que, por trabajo, ocio o interacción, somos dependientes de la IA. La creciente dependencia de la IA en diversas áreas de la vida plantea preocupaciones sobre la deshumanización de las interacciones y la pérdida de habilidades y capacidades humanas: si determinados roles, trabajos, acciones, actitudes, relaciones etc., van a ser sustituidas por máquinas, ¿quién asegura que, por desuso, no se pierdan? Y, si se pierden ¿qué clase de humano se obtendrá como consecuencia? ¿Estamos frente a la deshumanización del ser humano? El sintagma mismo es una aporía sin retorno: deshumanización de lo humano. Y vamos más allá: ¿cómo serán las máquinas desarrolladas por ese humano deshumanizado? Y más: ¿cómo terminarán siendo esos otros humanos que serán complementados y hasta suplantados por esas

máquinas? Hay una espiral ahí de deshumanización al que habría que prestarle muchísima atención.

18. Control y poder: Michel Foucault fue uno de los filósofos más importantes del poder y la biopolítica. Su concepto de un poder que se ejerce (no se tiene) y siempre "contra" algo/alguien como manera de entender todas las relaciones sociales, dio lugar a la transformación de la máxima "hacer morir-dejar vivir" que actuaba como ordenador de las sociedades antiguamente, por otra: "hacer vivir, dejar morir". Esta última frase que caracteriza la biopolítica da cuenta de un poder que atraviesa la vida toda. Si bien Foucault no pudo avizorar la IA y los cambios que ella introduciría en las relaciones de poder, hoy podríamos afirmar que ella es también un dispositivo biopolítico por medio del cual se hace vivir y se deja morir. La IA, como es de esperar, puede concentrar el poder en manos de unos pocos, lo que consecuentemente plantea preocupaciones sobre el control, la vigilancia y la manipulación, a través de su uso y desarrollo como dispositivos a la manera foucaultiana. Asimismo, plantea el interrogante acerca de las asimetrías que, posiblemente, se reproduzcan y profundicen: entre

países, entre bloques de países y, por supuesto, entre personas y grupos de personas.

19. Impacto medioambiental: no todo es beneficio en el desarrollo y uso de la IA: es hora de pensar desde una perspectiva de la casa común, la cual debemos cuidar y preservar pensando en un desarrollo sustentable. Las tecnologías y sus dispositivos pueden tener un impacto en el medio ambiente a causa del consumo de energía y de la generación de desechos electrónicos. En este sentido es imperativo reflexionar sobre la sostenibilidad y la responsabilidad ambiental en la implementación de la IA, pensando en el presente con los ojos puestos en el futuro.

> No queremos con este capítulo final contribuir la utopía efervescente y confiada, ni a la distopía desesperanzadora. Solo ponemos sobre la mesa en un plano rebatido algunas certezas y muchas incertidumbres.

20. Intimidad y vigilancia masiva: la IA se maneja con una cantidad enorme de datos personales que recopila, analiza y usa permanentemente. Es ingenuo pensar que la distopía pensada por George Orwell en 1984 no está plenamente en funcionamiento, mediante vigilancia y control con diversos

fines: algunos más inocuos (por ejemplo, vender consumidores), otros más dañinos, como manipular la opinión pública o someter a los pueblos. La privacidad pareciera, al día de hoy, solo un simulacro, y por eso es necesario crear conciencia y leyes que protejan las intimidades y nos protejan a los usuarios de quienes detentan el poder.

21. Impacto en el trabajo humano: la implementación de la IA en el lugar de trabajo plantea interrogantes y desafíos que ya está evaluando y denunciando la OIT respecto del empleo, la empleabilidad, la reestructuración de la economía y la necesidad de reentrenamiento y adaptación laboral. En un trabajo de agosto de 2023 concluye que hasta el día de hoy los empleos e industrias están parcialmente expuestos a la automatización de la IA Generativa, por lo cual no habría una pérdida de empleos sino una transformación para la que habrá que prepararse, y en este sentido la educación tiene un rol importante que desempeñar. El estudio es complejo y cruza datos a fin de profundizar en el análisis. De esta forma, también alerta que esa exposición será más importante en los países de renta alta que en los países de renta baja.

Asimismo, remarca también que los empleos más afectados serán los femeninos —agudizando la ya impuesta brecha de género— y los administrativos.

Si bien, entonces, la OIT no ve un peligro inminente en la IA incorporada al trabajo, es también importante poner este asunto en agenda, a fin de proceder a la preparación del escenario para que el futuro no nos sorprenda (Gmyrek, Berg, Bescond, 2023).

22. Equidad en la distribución de beneficios: la IA puede generar beneficios económicos y sociales significativos, pero también puede exacerbar las desigualdades existentes si los beneficios se concentran en manos de unos pocos, como todo bien transable en esta sociedad capitalista en la que vivimos. En este sentido, lo que se plantea es un desafío en la distribución de beneficios y la necesidad de políticas de inclusión y justicia social.

¿Y entonces?

No queremos con este capítulo final contribuir a la utopía efervescente y confiada, ni a la distopía desesperanzadora. Solo ponemos sobre la mesa en un

plano rebatido algunas certezas y muchas incertidumbres. No es malo descubrir que tenemos para ofrecer más interrogantes que respuestas: sabemos positivamente que no hay respuestas posibles si no se hacen las preguntas adecuadas y también sabemos, por experiencia y por método, que la desestabilización del *statu quo* a través de las preguntas es el primer paso de la reflexión y de la respuesta. Incluso cuando no hay respuestas.

La IA, como todo adelanto científico y técnico, aporta y aportará múltiples beneficios en todas las áreas del desarrollo humano. No se trata de dar la espalda al adelanto por temor a las consecuencias. De todos modos, el adelanto nos pasará por encima con y sin nuestro consentimiento y anuencia. Se trata de poner al ser humano en el centro a fin de empezar a construir una ética que dé marco a todos los planteamientos filosóficos que enlistamos precedentemente.

Seguramente habrá otros, la lista no pretende ser exhaustiva. La intención es abrir vías de análisis y no cerrar puertas: como todo adelanto científico y tecnológico, también, la IA desatará innumerables

peligros y “monstruos”. Creemos, sin embargo, que, si lo humano es puesto como catalizador de la experiencia en el desarrollo y uso de la superinteligencia, podemos esperar un futuro que no sea amenazador más allá del portal que estamos a punto de trasponer.

Mucho hemos escuchado que no se debe *tecnologizar* al ser humano, sino, por el contrario, debemos humanizar la tecnología. Claro está, que esta frase nos propone pensar más allá de una simple lectura y comprender la urgencia del análisis filosófico ante los avances tecnoemergentes. Con base en esto podemos preguntarnos ¿Hasta dónde podemos humanizar la tecnología y en especial a la inteligencia artificial? Para ello, no podemos dejar de preguntarnos ¿qué dimensiones implica el concepto de persona?, ¿qué implica ser humano? Ser persona, ser humano, en todo su espectro, abarca lo físico, lo afectivo, la voluntad, lo intelectual y lo espiritual. Por lo tanto, ¿hasta qué punto intentaremos y propondremos, como creadores de IA en su sentido de herramientas, dotarlas de estas dimensiones que nos hacen ser personas? Si pudiéramos lograr darles a las IA la dimensión

espiritual, hablaríamos entonces y tal vez de un horizonte tecno humano, en el que se entienda y se trabaje desde el concepto de complementariedad y no de sustitución.

El ser humano ha creado estas herramientas de IA muy sofisticadas en su concepción y diseño, pero ellas, las IA, solo nos ven a nosotros —las personas— como simples datos, números, que son la base de sus proyecciones para predecir posibles conductas y decisiones. Asimismo, estas tecnologías están cambiando el esquema de desarrollo en todas las esferas de la sociedad, desconociendo en sus formas algorítmicas la esencia de *ser persona*, reemplazando el crecimiento lineal por uno exponencial. También debemos en este punto preguntarnos quién es quién en este juego de desarrollo tecno humano ¿somos nosotros, las personas, los creadores de IA? ¿o las IA terminan condicionando todos nuestros deseos y decisiones en esta nueva era digital, con base a su ontología *cyborg*?

Elegimos terminar con esta sentencia: "La inteligencia artificial no es solo una tecnología, es una

forma de pensar. Y la forma en que pensamos sobre la inteligencia artificial da forma a lo que se convierte". (Turkle, 2015).

Referencias bibliográficas

Baudrillard, J. (1993). *Cultura y simulacro*. Kairós.

Baudrillard, J. (1996). *El crimen perfecto*. Anagrama.

Baudrillard, J. (1997). *La transparencia del mal*. Anagrama.

Bostrom, N. (2005). A History of Transhumanist Thought. *Journal of Evolution and Technology,* Vol.14, No. 1

Bostrom, N. (2016). Superinteligencia. Caminos, peligros, estrategias. Zaragoza: Teell Editorial

Bostrom, N., (2002). Existential Risks: analyzing human extinction scenarios and related hazards. *Journal of Evolution and Technology*, vol. 9, n.o 1.

Chomsky, N. (1992). *El control de los medios de comunicación*. Recuperado de: https://cronicon.net/paginas/Documentos/paq2/No.31.pdf

Danaher, J. (2016) "The Threat of Algocracy: Reality, Resistance and Accommodation". *Philosophy & Technology*. Volume 29, 3, p. 245–268. https://doi.org/10.1007/s13347-015-0211-1. Recuperado de: https://www.researchgate.net/publication/291015337_The_Threat_of_Algocracy_Reality_Resistance_and_Accommodation

Debord, G. (1967). *La sociedad del espectáculo*. Recuperado de: http://serbal.pntic.mec.es/~cmunoz11/Societe.pdf

FHEM8239

Foucault, M. (1979). *Microfísica del poder*. Madrid: Ediciones de la Piqueta. Recuperado de: https://www.pensamientopenal.com.ar/system/files/2014/12/doctrina39453.pdf

Foucault, M. (s/f). *Power*. Recuperado de: https://monoskop.org/images/b/b9/Foucault_Michel_Power_2000.pdf

Gmyrek, P., Berg, J., Bescond, D. 2023. Generative AI and jobs: A global analysis of potential effects on job quantity and quality, *ILO Working Paper 96* (Geneva, ILO). https://doi.org/10.54394/

Gracia, D. (2009). *La deliberación moral. el papel de las metodologías en ética clínica*. Universidad Complutense

Harari, Y (2017). *Homo Deus: Breve historia del mañana*. Debate

Han, B. (2014). *En el enjambre*. Herder

Han, B. (2022). *Infocracia*. Penguin Random House

Lasswell, H. (1927). Theory of Political Propaganda. The American Political Science

Lasswell, H. (1935). *World Politics and Personal Insecurity*. McGraw Hill.

McLuhann, M. (1996). *Comprender los medios de comunicación. Las extensiones del ser humano.* Paidós.

Morte Ferrer, R., Monasterio Astobiza, A., Ausín, T. (2017). Presentación. Ética de datos, sociedad y ciudadanía. *Dilemata*. Núm. 24 (2017): Recuperado de: https://www.dilemata.net/revista/index.php/dilemata/article/view/412000116

Padilla, E. (2023). *Entre la celebración y el rechazo -La incidencia de la tecnociencia en la conceptualización de lo humano*. Prometeo.

Rainie, L., Wellman, B. (2012). Networked: The New Social Operating System. *The MIT Press Review*, 21(3), 627-631. Recuperado de: https://www.cambridge.org/core/journals/american

-political-science-review/article/abs/theory-of-political-propaganda/9AFAA9A4B7BD71EA20B3521A99D7EF3F

Searle, J. (2000). *Razones para actuar: Una teoría del libre albedrío*. Oviedo: Nobel.

Searle, John. R. (1980) Minds, brains, and programs. Behavioral and Brain. *Sciences* 3 (3): 417-457. Recuperado de: https://web-archive.southampton.ac.uk/cogprints.org/7150/1/10.1.1.83.5248.pdf

Turkle, S. (2015). *Reclaiming conversation: The power of talk in a digital age.* Penguin Books.

Valente, L. (2019). La persona electrónica. *Revista Anales de la Facultad de Ciencias Jurídicas y Sociales*. Universidad Nacional de La Plata. UNLP. Año 16/Nº 49-2019. Anual. Impresa ISSN 0075-7411-Electrónica ISSN 2591-6386. Recuperado de: file:///D:/Users/W10/Downloads/lucianaali,+1Luis+Alberto+Valente+(Anales+49).pdf

Sobre el autor

Dr. Gustavo de Elorza Feldborg
PhD. en Tecnología Educativa

Experto en Neuroeducación y Aprendizaje en la Era Digital
El Dr. Gustavo de Elorza Feldborg es un reconocido *Doctor en Tecnología Educativa* con amplia experiencia en investigación y docencia a nivel universitario y de posgrado internacional. Su trayectoria se destaca por su enfoque innovador en la *intersección entre la neurociencia, el aprendizaje y las tecnologías emergentes,* particularmente en el ámbito de la educación virtual e inmersiva.

Investigador y Líder en Educación del Siglo XXI
El Dr. de Elorza Feldborg posee un *Posdoctorado en Educación* en el que profundizó en el estudio de la "Neurociencia, Cerebro y Redes de Aprendizaje en Educación Virtual". Además, cuenta con la prestigiosa certificación como *Líder de Aprendizaje* otorgada

por la Universidad de Harvard, lo que ratifica su compromiso con la vanguardia en el campo educativo.

Conferencista TEDx y escritor prolífico

Su pasión por la educación lo ha llevado a compartir sus conocimientos en diversos escenarios, incluyendo conferencias TEDx y publicaciones como autor. Su capacidad para comunicar ideas complejas de manera clara y *engaging* lo convierte en un referente en el ámbito de la neuroeducación y el aprendizaje en la era digital.

A la vanguardia de la innovación educativa

En la actualidad, el Dr. de Elorza Feldborg se encuentra trabajando en el desarrollo de un *modelo pedagógico innovador* que integra los últimos avances en neuroeducación, metaverso, tecnologías digitales e inteligencia artificial, con el objetivo de transformar la educación y potenciar el aprendizaje en un mundo cada vez más conectado.

Experiencia internacional y trayectoria destacada

Su labor ha impactado a miles de personas en todo el mundo a través de conferencias, cursos y talleres dictados en España, México, Colombia, Uruguay, Bolivia, Perú, Panamá, Chile, EE.UU. y Argentina. El Dr. de Elorza Feldborg se consolida como un referente indiscutible en el campo de la educación del siglo XXI, abriendo camino hacia un futuro en el que el aprendizaje sea más efectivo, personalizado, significativo y comprensivo para todos.

Otros libros del autor

Los seres humanos tenemos la capacidad de pensar de manera abstracta, hacer conexiones creativas, resolver problemas complejos y generar nuevas ideas e innovaciones. En este sentido, la IA viene logrando avances en la generación de contenido y en tareas creativas, aunque por el momento todavía está limitada en términos de originalidad y pensamiento en comparación con los seres humanos.

Es decir, aunque la inteligencia artificial ha avanzado significativamente en muchas áreas y puede superar a los humanos en tareas específicas, tanto en sus dimensiones de velocidad como en su capacidad de procesamiento, todavía está lejos de replicar la complejidad y diversidad de la inteligencia humana en su conjunto. La inteligencia artificial actual carece de la experiencia, de conciencia, de emociones, de creatividad y de la comprensión profunda que caracterizan a la inteligencia humana.

Humanizar la tecnología parece un sinsentido, y mucho más, intentar lograr una tecnología del alma, sobre todo para aquellos que pretenden tecnificar la humanidad, mejorarla, trascenderla por medio de la IA, de los metaversos, y de toda la gama de tecno herramientas que tenemos al presente. Y, sin embargo, es un horizonte que no debemos ignorar quienes trabajamos en educación, puesto que la intersección entre la educación, el metaverso y la IA —ya presente y en acción— no solo representa un desafío, sino también una oportunidad sin precedentes para redefinir el paradigma educativo y preparar a las generaciones futuras para un mundo en constante evolución.

En **El despertar de las máquinas pensantes**, el Dr. de Elorza Feldborg pone el acento no solo en los aspectos técnicos de ese maridaje entre educación e IA, sino en todos los interrogantes éticos que están involucrados y no pueden ser desestimados:
tal es la seriedad de la cuestión.